

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

*OM protein - Protein search, using SW model
 Run on: June 23, 2005, 08:32:43 ; Search time 42.1481 Seconds
 (without alignments)
 3578.727 Million cell updates/sec

Title: US-10-697-828-8
 Perfect score: 2051
 Sequence: 1 MWLPRFSSKTVTLLIAQTT.....LTLDVLPRGPDHFSWASPD 390

Scoring table: BLOSUM62
 Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
 Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 10%
 Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: GeneseqP1980s:*
- 2: GeneseqP1990s:*
- 3: GeneseqP2000s:*
- 4: GeneseqP2001s:*
- 5: GeneseqP2002s:*
- 6: GeneseqP2003as:*
- 7: GeneseqP2003bs:*
- 8: GeneseqP2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Match Length | DB ID | Description | OS | Homo sapiens. |
|------------|--------|-------|--------------|------------|---|----|--|
| 1 | 2051 | 100.0 | 390 | 4 AAY72639 | Aay72639 Human gly Abb81556 Human int Abb41947 Human ORF | PN | W0200106015-A1. |
| 2 | 2051 | 100.0 | 390 | 5 AAB81556 | Abb81555 Consensus | XX | |
| 3 | 1865.5 | 91.0 | 418 | 3 AAB41947 | Aay72640 Human gly Abb81554 Human cor Aae15438 Human dru Adi21086 Novel hum Adl61235 Human tyr | PD | 25-JAN-2001. |
| 4 | 1758.5 | 85.7 | 395 | 5 ABB81555 | Aay72640 Human gly Abb81554 Human cor Aae15438 Human dru Adi21086 Novel hum Adl61235 Human tyr | XX | 19-JUL-2000; 2000WO-US019741. |
| 5 | 1729.5 | 84.3 | 395 | 4 AAY72640 | Aay72640 Human gly Abb81554 Human cor Aae15438 Human dru Adi21086 Novel hum Adl61235 Human tyr | XX | 20-JUL-1999; 99US-0144694P. |
| 6 | 1729.5 | 84.3 | 395 | 5 AAE15438 | Aay72638 Mouse gly Aau11275 Murine int Abb81557 Mouse int Adj70405 Human hea Aam93309 Human pol Adl3784 Human pro | PR | 13-JUN-2000; 2000US-00593828. |
| 7 | 1729.5 | 84.3 | 395 | 5 AAE15438 | Aay72638 Mouse gly Aau11275 Murine int Abb81557 Mouse int Adj70405 Human hea Aam93309 Human pol Adl3784 Human pro | XX | WPI; 2001-138471/14. |
| 8 | 1729.5 | 84.3 | 395 | 5 AAE15438 | Aay72638 Mouse gly Aau11275 Murine int Abb81557 Mouse int Adj70405 Human hea Aam93309 Human pol Adl3784 Human pro | DR | N-PDBB; AAD02597, AAD02699. |
| 9 | 1729.5 | 84.3 | 395 | 8 ADL61235 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor Abb81558 Human cor | XX | New Glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications. |
| 10 | 1536.5 | 74.9 | 395 | 4 AAY72638 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | XX | The present sequence is human glycosyl sulfotransferase-4alpha (GST-4 alpha). GST-4 gene is found on chromosome 16q23.1. GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non |
| 11 | 1536.5 | 74.9 | 395 | 5 AAB81557 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | CC | CC |
| 12 | 1523 | 74.3 | 418 | 5 ABB81557 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | PT | CC |
| 13 | 1513.5 | 73.8 | 394 | 7 ADJ70405 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | PT | CC |
| 14 | 1013 | 49.4 | 386 | 8 AAM93309 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | PS | Claim 3; Fig 1; 128pp; English. |
| 15 | 1013 | 49.4 | 386 | 8 AAM93309 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | XX | |
| 16 | 1008 | 49.1 | 386 | 2 AAY39918 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | XX | |
| 17 | 1003.5 | 48.9 | 380 | 5 ABB81557 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | PT | |
| 18 | 970 | 47.3 | 386 | 3 AAY72619 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | PT | |
| 19 | 958.5 | 46.7 | 388 | 2 AAY39919 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | XX | |
| 20 | 808 | 39.4 | 169 | 5 ABB81559 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | XX | |
| 21 | 755 | 36.8 | 169 | 5 ABB81559 | Aay39918 Human gly Aau11274 Human L-s Aay72619 Human tra Aay39919 Mouse gly Abb81559 Human cor | XX | |
| 22 | 665 | 32.4 | 483 | 2 AYX31656 | Aay31656 Mouse N-a Abb93367 Human pro | CC | |
| 23 | 659.5 | 32.2 | 530 | 5 AAB93367 | Adg1590 Human sof | CC | |
| 24 | 659.5 | 32.2 | 530 | 8 ADQ18590 | Aay31657 Human N-a | CC | |
| 25 | 656.5 | 32.0 | 484 | 2 AYX31657 | | CC | |

ALIGNMENTS

RESULT 1

AAV72639 standard; protein; 390 AA.

ID AAV72639;

AC AAY72639;

XX

DT 02-MAY-2001 (first entry)

DE Human glycosyl sulfotransf erase-4alpha (GST-4alpha).

KW Human; glycosyl sulfotransf erase-4alpha; GST-4alpha; immunosuppressive;

KW therapy; selectin binding inhibitor; gene therapy; inflammation;

KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;

KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;

KW Glomerulonephritis; myasthenia gravis; Sjogren's syndrome; Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;

KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;

KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;

KW asthma; hypersensitivity; rheumatic fever; tissue rejection;

KW chromosome 16q23.1.

XX Homo sapiens.

XX

-sulphated selectin ligand, GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in inhibiting a selectin mediated binding event. GST is useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis, myasthenia gravis, Sjögren's syndrome, Hashimoto's disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious anaemia, demyelinating diseases, Cirrhosis, ulcerative colitis, dermatitis, myocarditis, regional enteritis, adult respiratory distress syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation. Note: The present sequence is also shown in sequence listing (page no: 56) but lacks four nucleotides at its 3' end

Sequence 390 AA;

Query Match 100.0%; Score 2051; DB 4; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.2e-215;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Db 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Qy 61 FSOHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120
Db 61 FSOHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120

Qy 121 FFWNTSRALCSPACAFPRGTISKODVCKTLCTROFSLAREACRYSVHVKVEFRFF 180
Db 121 FFWNTSRALCSPACAFPRGTISKODVCKTLCTROFSLAREACRYSVHVKVEFRFF 180

Qy 181 NLQVLYPLISDPALNRLRVHLYDPAVLRSREAAAGPILARONGIVLIGTNGKVEADPHL 240
Db 181 NLQVLYPLISDPALNRLRVHLYDPAVLRSREAAAGPILARONGIVLIGTNGKVEADPHL 240
Qy 241 RLIREVCVRSHVRIAEAATLKKPPPLGRYLRVRFEDLAREPLAEIRALYAFGLTLPQL 300
Db 241 RLIREVCVRSHVRIAEAATLKKPPPLGRYLRVRFEDLAREPLAEIRALYAFGLTLPQL 300

Qy 301 EAINTNITHGSGIGKPIEAFHTSSRNARNVSQAWRHLPFTKILRYOEVCAAGALQIGYR 360
Db 301 EAINTNITHGSGIGKPIEAFHTSSRNARNVSQAWRHLPFTKILRYOEVCAAGALQIGYR 360

Qy 361 PIVYSDAQORDLTLDVLVRGPDHFSWASPD 390
Db 361 PIVYSDAQORDLTLDVLVRGPDHFSWASPD 390

RESULT 2

ABB81556 standard; protein; 390 AA.

XX ABB81556; Human intestinal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:4.
AC ABB81556; Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;
KW corneal; sulfate; keratan sulfate; macular corneal dystrophy; MCD;
KW ophthalmological.

DT 05-SEP-2002 (first entry)

XX Homo sapiens.
OS Homo sapiens.
XX US2002061562-A1.
PN US2002061562-A1.
PD 23-MAY-2002.

XX 09-AUG-2001; 2001US-00927602.

XX 11-AUG-2000; 2000US-00638211.

PR PR

PR 11-AUG-2000; 2000US-0325773P.

XX (FUKUJ/ FUKUDA M. N.
PA (AKAM/ AKAMA T. O.XX Fukuda MN,
PI Akama TO;
XX DR WPI: 2002-507643/54.

XX New nucleic acid encoding corneal N-acetylglucosamine-6-sulfotransferase, useful for treatment, monitoring and diagnosis of macular corneal dystrophy.
XX Example 5; Fig 2A-B; 69pp; English.

XX The present invention describes human corneal N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse sulfation of keratan sulfate (KS). Also described is a method for monitoring the effect of treatments for macular corneal dystrophy (MCD), and detecting susceptibility to MCD. (I) is located to chromosome 16q22, and has ophthalmological activity. (I) can be used to treat or prevent macular corneal dystrophy types I or II. (I) makes possible treatment of MCD without requiring keratoplasty or keratectomy. The present sequence represents human intestinal N-acetylglucosamine-6-sulfotransferase, which is given in comparison with (I) in the exemplification of the present invention

XX Sequence 390 AA;

Query Match 100.0%; Score 2051; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.2e-215;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Db 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Qy 61 FSQHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120
Db 61 FSQHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120

SQ Sequence 390 AA;

Query Match 100.0%; Score 2051; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.2e-215;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Db 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Qy 61 FSQHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120
Db 61 FSQHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120

SQ Sequence 390 AA;

Query Match 100.0%; Score 2051; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.2e-215;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Db 1 MWLPRFSSKTVVLLAQTCLLFLIISRPGPSPAGEDRVVLVLSWRGSSPFSQL 60
Qy 61 FSQHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120
Db 61 FSQHPDVYLMEPAWHVVWTTLSQSGAATLHAYVTLDMRSIFLCMDMVDAYMPQRNLSA 120
Qy 121 FENWATSRALCSPACAFPRGTISKODVCKTLCTROFSLAREACRYSVHVKVEFRFF 180
Db 121 FENWATSRALCSPACAFPRGTISKODVCKTLCTROFSLAREACRYSVHVKVEFRFF 180

SQ Sequence 390 AA;

Query Match 100.0%; Score 2051; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.2e-215;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 FENWATSRALCSPACAFPRGTISKODVCKTLCTROFSLAREACRYSVHVKVEFRFF 180
Db 121 FENWATSRALCSPACAFPRGTISKODVCKTLCTROFSLAREACRYSVHVKVEFRFF 180
Qy 181 NLQVLYPLISDPALNRLRVHLYDPAVLRSREAAAGPILARONGIVLIGTNGKVEADPHL 240
Db 181 NLQVLYPLISDPALNRLRVHLYDPAVLRSREAAAGPILARONGIVLIGTNGKVEADPHL 240
Qy 241 RLIREVCVRSHVRIAEAATLKKPPPLGRYLRVRFEDLAREPLAEIRALYAFGLTLPQL 300
Db 241 RLIREVCVRSHVRIAEAATLKKPPPLGRYLRVRFEDLAREPLAEIRALYAFGLTLPQL 300
Qy 301 EAINTNITHGSGIGKPIEAFHTSSRNARNVSQAWRHLPFTKILRYOEVCAAGALQIGYR 360
Db 301 EAINTNITHGSGIGKPIEAFHTSSRNARNVSQAWRHLPFTKILRYOEVCAAGALQIGYR 360
Qy 361 PIVYSDAQORDLTLDVLVRGPDHFSWASPD 390
Db 361 PIVYSDAQORDLTLDVLVRGPDHFSWASPD 390

RESULT 3

ABB41947 ID AAB41947 standard; protein; 418 AA.

XX AAB41947; Human ORF1711 polypeptide sequence SEQ ID NO:3422.
AC AAB41947; Human ORF1711 polypeptide sequence SEQ ID NO:3422.
XX DE 08-FEB-2001 (first entry)
XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;

| | | | |
|--------|---|----|---|
| KW | vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; | Db | 88 FSQHPDPFYLMEPAWHWTTLSQGSAATLHMAYRDLIERSIFLCMDIVFDAYMPEGPRROS |
| KW | anticonvulsant; oseopathic; antiarthritic; immunosuppressant; cardiotonic; | Qy | 120 AFFNWATSRALCSPPACSAFPRTTSKQDVCKTLCTRQPSLAREACSYSHVVLKEVRF |
| KW | immunomodulant; thrombolytic; coagulant; vasotropics; antidiabetic; | Db | 148 SLFOWENSRALCSAPCDIIPDEIIPIRAHCRLLCSQQPPEVKACRSYSHVVLKEVRF |
| KW | hypotensive; dermatological; immunosuppressive; antiinflammatory; | Qy | 180 ENIQVLYPLISDPAFLNLRVHLVRDPAVRLRSREAAAGPILARDNGIVLTGNGKNEADPH |
| KW | antiviral; antibacterial; antifungal; antirheumatic; antithyroid; | Db | 208 ENIQSLYPLIKDPLNLNLRVHLVRDPAVRLRSREAAAGPILARDNGIVLTGNGKNEADPH |
| KW | antiaemic; gene therapy; cancer; proliferative disorder; hypertension; | Qy | 240 LRLIREVCRSHVRIAATLKPPPLGRYLVRFEDLAREPLAEIRALYAFTGLTLPQ |
| KW | neurodegenerative disorder; osteoarthritis; graft vs host disease; | Db | 268 LRLIREVCRSHVRIAATLKPPPLKDPDPLNLNLRVHLVRDPAVRLRSREAAAGPILARDNGIVLTGNGKNEADPH |
| KW | cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; | Qy | 300 LEAMINITHGSGIGKPIEAFHTSSRNARNVSQAWHALPFTKLRLRQEVCAGAQLLGY |
| KW | cholesterol ester storage; systemic lupus erythematosus; infection; | Db | 328 LEAMINITHGSGIGKPIEAFHTSSRNARNVSQAWHALPFTKLRLRQEVCAGAQLLGY |
| KW | severe combined immunodeficiency; malaria; autoimmune disorder; asthma; | Qy | 360 RPVYSSADQDQLTLIDLVLPRGDHFDSWASPD |
| KW | allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; | Db | 388 RPVYSSADQDQLTLIDLVLPRGDHFDSWASPD |
| KW | bone damage; cartilage damage; antinflammatory disease; coagulation; | XX | 418 |
| KW | thrombosis; contraceptive. | XX | |
| OS | Homo sapiens. | XX | |
| XX | WO2005B873-A2. | XX | |
| PN | | XX | |
| PD | 05-OCT-2000. | XX | |
| XX | | XX | |
| PF | 31-MAR-2000: 20000HQ-US008621. | XX | |
| XX | | XX | |
| PR | 31-MAR-1999; 9905-0127607P | XX | |
| PR | 02-APR-1999; 9905-0127636P | XX | |
| PR | 05-APR-1999; 9905-0127728P | XX | |
| PR | 30-MAR-2000: 20000US-00540763. | XX | |
| XX | | XX | |
| PA | (CURA-) CURAGEN CORP. | XX | |
| XX | | XX | |
| PI | Shimkets RA, Leach M; | XX | |
| XX | | XX | |
| DR | WPI: 2000-602362/57. | XX | |
| N-PDBB | AAC76156. | XX | |
| XX | | XX | |
| PA | Novel nucleic acids and peptides derived from open reading frame X, | XX | |
| XX | useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease. | XX | |
| XX | | XX | |
| PS | Claim 11: Page 2599-2600: 5507PP: English. | OS | |
| XX | | OS | |
| CC | AAC7446 to AAC77606 encode the proteins given in AAB43397 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotoxic; vulnerability; antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunomimulant; cardiotonic; thrombolytic; coagulant; vasotropics; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antithyroid; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianametic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathologic conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disease, disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antihypertensive disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive. | XX | |
| CC | Sequence 418 AA; | XX | |
| CC | Query Match Score 1865.5; DB 3; Length 418; | XX | |
| CC | Best Local Similarity 91.0%; Pred. No. 3.1e-195; | FT | Misc-difference 380 |
| CC | Matches 356; Conservative 9; Mismatches 25; Indels 1; Gaps 1; | FT | Misc-difference 382 |
| CC | 1 MWLPRSSKTVVLLAQTCLLFLISRPSSPAGGEDEVHVLSSRSGSSFLGQI 60 | FT | Misc-difference 384 |
| CC | 28 MWLPRSSKTVVLLAQTCLLFLISRPSSPAGGEDEVHVLSSRSGSSFLGQI 87 | FT | Misc-difference 390 |
| CC | 61 FSQHPDPFYLMEPAWHWTTLSQGSAATLHMAYRDLIERSIFLCMDIVFDAYM-PQRNL 119 | FT | Misc-difference 391 |

| | | | | | | | |
|----------|---|---|-----|----------|--|---|--|
| Db | 1 | MWLPRVSSTAVTALLAQ-TELLFLVSRGPSSPAGGEARVHYVLSSWRSGSSFYVQL | 59 | | CC | without requiring keratinoplasty or keratectomy | |
| Oy | 61 | FSDQHDVFYIPLMEPAWHVNTTSLSGSAATLHMAYRDLMRSTFLCDMDVFDAYMPOSRLNSA | 120 | XX | Sequence 395 AA; | | |
| Db | 60 | FNQHPDVFTLMEPAWHVNTTSLSGSAATLHMAYRDLVISFLCDMDVFDAYLPWRRNLSD | 119 | QQ | Query Match Score 1729.5; DB 5; Length 395; | | |
| Qy | 121 | FENPRTSRLCSPPACSAFPRTGKTKDVKCPLCTROPEFLAREACRSYSHVLTKEVRF | 180 | QQ | Best Local Similarity 85.8%; Prod. No. 2.e-180; | | |
| Db | 120 | LFWAVASRLCSPPACSAFPRTGKTKDVKCPLCARQSTLAREACRSYSHVLTKEVRF | 179 | Db | Matches 333; Conservative 17; Mismatches 37; Indels 1; | | |
| Qy | 181 | NLQVLYPLISDPALNRLTVLVRDPRAVLRSREAGPLIARDNGIVLGTNKTVTEADPHL | 240 | Qy | 1 MWLPRVSSTAVTALLAQ-TFLLELISRPGSFSPAGEARVHYVLSSWRSGSSFYVQL | 60 | |
| Db | 180 | NLQVLYPLISDPALNRLTVLVRDPRAVLRSREQTAKALARNDGIVLGTNTGTVTEADPHL | 239 | Db | 1 MWLPRVSSTAVTALLAQ-TFLLELISRPGSFSPAGEARVHYVLSSWRSGSSFYVQL | 59 | |
| Qy | 241 | RLETRVCRSHVRIAEEATLKPPLFLGRYRVLVRFEDLAREPLAIRALYAFGLTLTPQL | 300 | Qy | 61 FSCOPHDVYLMPEAWHWTTLSGSAATLHMAYRDLMRSTFLCDMDVFDAYLPWRRNLSD | 120 | |
| Db | 240 | RVREVCRSHVRIAEEATLKPPLFLGRYRVLVRFEDLAREPLAIRALYAFGLTLTPQL | 299 | Qy | 121 FFWATSRALCSPPACSAFPRTGKTDVKCPLCTROPEFLAREACRSYSHVLTKEVRF | 180 | |
| Qy | 301 | EAWHNITHGIGKPIEAFHTSSRNARNVSQMRHALPTFKILRQEVCAAGALQLGXR | 360 | Db | 120 LFQAVASRLCSPPACSAFPRTGKTKDVKCPLCARQSTLAREACRSYSHVLTKEVRF | 179 | |
| Db | 300 | EAWHNITHGIGKPIEAFHTSSRNARNVSQMRHALPTFKILRQEVCAAGALQLGXR | 359 | Qy | 181 NLQVLYPLISDPALNRLTVLVRDPRAVLRSREAGPLIARDNGIVLGTNKTVTEADPHL | 240 | |
| Qy | 361 | PVYSAQQDQLTLDDLVLPDPDHFSWAS | 388 | Qy | 180 NLQVLYPLISDPALNRLTVLVRDPRAVLRSREAGPLIARDNGIVLGTNKTVTEADPHL | 239 | |
| Db | 360 | PVYSEDEQRNLALDDLVLPQLNGFTWAS | 387 | Db | 180 RLIREVCRSHVRIAEEATLKPPLFLGRYRVLVRFEDLAREPLAIRALYAFGLTLTPQL | 299 | |
| <hr/> | | | | | | | |
| RESULT 6 | | | | | | | |
| ABB81554 | ID | EATHNITHGIGKPIEAFHTSSRNARNVSQMRHALPTFKILRQEVCAAGALQLGXR | 360 | Qy | 301 EAHTNITHGIGKPIEAFHTSSRNARNVSQMRHALPTFKILRQEVCAAGALQLGXR | 360 | |
| ABB81554 | XX | standard; protein; 395 AA. | | Db | 300 EAHTNITHGIGKPIEAFHTSSRNARNVSQMRHALPTFKILRQEVCAAGALQLGXR | 359 | |
| AC | ABB81554; | | | Qy | 361 PVYSAQQDQLTLDDLVLPDPDHFSWAS | 388 | |
| XX | DT | 05-SEP-2002 (first entry) | | Db | 360 PVYSEDEQRNLALDDLVLPQLNGFTWAS | 387 | |
| DE | Human corneal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:2. | | | <hr/> | | | |
| XX | KW | Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST; | | RESULT 7 | | | |
| KW | KW | corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD; | | ID | AAB15438 standard; protein; 395 AA. | | |
| KW | KW | ophthalmological; chromosome 16q22. | | XX | AAB15438; | | |
| OS | Homo Sapiens. | | | AC | | | |
| XX | PN | US2002061562-A1. | | XX | 12-MAR-2002 (first entry) | | |
| PD | XX | 23-MAY-2002. | | XX | Human drug metabolising enzyme (DME)-5. | | |
| XX | XX | | | XX | Human; drug metabolising enzyme; gene therapy; autoimmune disorder; | | |
| PF | XX | 09-AUG-2001; 2001US-00927602. | | KW | acquired immune deficiency syndrome; infection; | | |
| XX | XX | 11-AUG-2000; 2000US-00628211. | | KW | inflammatory disorder; adult respiratory distress syndrome; cell proliferative disorder; | | |
| PR | PR | 11-AUG-2000; 2000US-0032573P. | | KW | AIDS; conjunctivitis; actinic keratosis; arteriosclerosis; | | |
| XX | XX | (FUKU/ FUKUDA M. N. PAKAM/ AKAMA T. O. | | KW | cancer; endocrine disorder; hypothalamus disorder; pituitary disorder; gastrointestinal disorder; metabolic disorder; developmental disorder; liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis; goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder; | | |
| PA | XX | | | KW | DME-5. | | |
| PI | XX | Fukuda MN, Akama TO. | | XX | Homo sapiens. | | |
| DR | XX | WPI; 2002-507643/54. | | OS | | | |
| DR | XX | N-PSDB; ABN89506. | | XX | Location/Qualifiers | | |
| PT | XX | New nucleic acid encoding corneal N-acetylglucosamine-6-sulfotransferase, useful for treatment, monitoring and diagnosis of macular corneal dystrophy. | | Key | 1-35 | | |
| PT | XX | The present sequence represents human corneal N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse sulfation of keratan sulfate (KS). Also described is a method for monitoring the effect of treatments for macular corneal dystrophy (MCD), and detecting susceptibility to MCD. (I) is located on chromosome 16q22, and has ophthalmological activity. (I) can be used to treat or prevent macular corneal dystrophy types I or II. (I) makes possible treatment of MCD | | Peptide | /label= Signal_Peptide | | |
| PS | XX | Claim 13; Fig 1A-D; 69pp; English. | | FT | 1-32 | | |
| PS | XX | | | Peptide | | | |
| PT | XX | | | Protein | 33...335 | | |
| PT | XX | | | Protein | /note= "Human mature DME-5 protein" | | |
| PT | XX | | | Protein | 36...395 | | |
| PT | XX | | | Protein | /note= "Human mature DME-5 protein" | | |
| XX | XX | | | XX | WO200173468-A2. | | |
| XX | XX | | | XX | 25-OCT-2001. | | |
| CC | CC | | | XX | XX | | |

| | | | | |
|-----------------------|---|----|---|-----|
| PF | 12-APR-2001; 2001WO-US011869. | Db | 180 NLQVLYPLLSDBALNRLIVHLVRDPRAVLRSREQTAKALARNDNGIVLTGNTWADPGL | 239 |
| XX | 13-APR-2000; 2000US-0197590P. | Qy | 241 RLJRECRSHYRIAATLKKPPFLGRYRFLYRFEDAREPLAEITALYAFTGLTLTPOL | 300 |
| PR | 19-APR-2000; 2000US-0198403P. | Db | 240 RVRECRSHYRIAATLKKPPFLGRYRFLYRFEDAREPLAEITALYAFTGLSLTPOL | 299 |
| PR | 20-APR-2000; 2000US-0200185P. | Qy | 301 EAWHNNTHGSSIGKPEAFTSSRNARNSQAWHLPFKILRYOEVCAQALQLGYR | 360 |
| PR | 05-MAY-2000; 2000US-0202234P. | Db | 300 EAWHNNTHGSGPAREFAEFTSSRNALNTSQAWHLPFKILRYOEVCAQALQLGYR | 359 |
| PR | 11-MAY-2000; 2000US-0203509P. | Qy | 361 PYYSADQDQDLVLPRGDHFMSAS | 388 |
| XX | (INCY-) INCYTE GENOMICS INC. | Db | 360 PYYSDEBORNALDLVLPRGLNGFTMAS | 387 |
| XX | POLICKY JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM; | Qy | | |
| PI | Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS; | Db | | |
| PI | Gandhi AR, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA; | | | |
| PI | Pi-Au-Young J; | | | |
| XX | WPI: 2002-066363/09. | | | |
| DR | N-PSDB; AAD24670. | | | |
| XX | Novel isolated human drug metabolizing enzymes referred as DME 1-10, useful for diagnosing, treating, or preventing disorders associated with aberrant expression of DME such as allergy, anemia, asthma, infertility. | | | |
| XX | Claim 1a; Page 131-132; 143PP; English. | | | |
| PS | The invention relates to human drug metabolising enzymes referred as DME and nucleic acid molecules encoding such enzymes. Polynucleotides of the invention are useful for assessing toxicity of test compounds and in gene therapy. Sequences of the invention are useful in the diagnosis, prevention and treatment of autoimmune/inflammatory disorders such as acquired immune deficiency syndrome (AIDS), adult respiratory distress syndrome, allergies, anaemia, atherosclerosis, asthma, autoimmune haemolytic anaemia, Goodpasture's syndrome, Crohn's disease, glomerulonephritis, Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis, osteoarthritis, osteoporosis, psoriasis, systemic lupus erythematosus, rheumatoid arthritis, scleroderma, ulcerative colitis, uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic infections; cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis, Duchenne and Becker dystrophy, hepatitis, Cushing's syndrome, cancers, myelodysplastic syndrome, epilepsy; endocrine disorders such as primary brain tumours, adenomas, infarction associated with pregnancy, aneurysms, vascular malformations; eye disorders such as conjunctivitis, iritis, retinitis, glaucoma, pigmentosa; metabolic disorders such as Addison's disease, cystic fibrosis, diabetes, goitre, glycogen storage diseases, hypercholesterolaemia, hyperthyroidism, hypoglycaemia, lipid myopathies, Menkes syndrome, mannosidosis, obesity; gastrointestinal disorders such as dysphagia, gastric carcinoma, anorexia, nausea, gastroenteritis, hyperbilirubinaemia, emesis, cirrhosis, diarrhoea, jaundice, Reye's syndrome, peliosis hepatitis, hepatic vein thrombosis and developmental disorders. The present sequence is human DME-5 protein | | | |
| XX | Sequence 395 AA; | | | |
| Query Match | 84.3%; Score 1729.5; DB 5; Length 395; | | | |
| Best Local Similarity | 85.8%; Prod. No. 2.4e-180; | | | |
| Matches | 333; Conservative 17; Mismatches 37; Indels 1; Gaps 1; | | | |
| Qy | 1 MWLPRFSSKTVTLLAQTCLLFLIISRPGPSPAGGEDRYHVLYLSSWSRGSSFLQQL | CC | The invention relates to an isolated polynucleotide encoding a polypeptide with biological activity. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders and other traits, to assess biodiversity, as nutritional sources or supplements. | |
| Db | 1 MWLPRVSSTAVTLLAQ-TFLFLFLVSRGPSPAGEARVHVLYLSSWSRGSSFVGQL | CC | The polynucleotides may also be used as molecular weight markers, chromosomal markers or map related gene positions, or as an antigen to raise anti-DNA antibodies or elicit immune response. The polypeptides are useful for raising antibodies, as markers for tissues in which the corresponding polypeptide is expressed, for re-engineering damaged or diseased tissue, for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and/or nerve tissue growth or regeneration, in wound healing, in tissue repair and replacement, in healing of burns, incisions and ulcers, and in treating cancer. The present sequence represents the amino acid sequence of a novel human | |
| Qy | 61 FSGHPDVYLMEPAWHWTM1SOGSAATTLMARDLMSIFLCMDYFDAYMPOSNLSA | CC | | |
| Db | 60 FNGHPDVYLMEPAWHWTM1SOGSAATTLMARDLMSIFLCMDYFDAYLPWNRLSD | CC | | |
| Qy | 121 FFWTWSRALCSPACSAFPAGTISKQDVCKLCTROPSLAREACSYSHVLYKEVRF | CC | | |
| Db | 120 LFQWAVSRLCSPACSAFPAGTISSEACKPKLARDSFTLAREACSYSHVLYKEVRF | CC | | |
| Qy | 181 NLQVLYPLLSDPALNRLIVHLVRDPRAVLRSREQTAKALARNDNGIVLTGNTWADPHL | CC | | |

| | | | |
|-----------|---|--|---|
| CC | protein. | PT | treating breast cancer. |
| XX | Sequence 395 AA; | XX | Claim 9; SEQ ID NO 159; 649PP; English. |
| SQ | Query Match 84.3%; Score 1729.5; DB 7; Length 395; Best Local Similarity 83.8%; Pred. No. 2.4e-180; Matches 333; Conservative 17; Mismatches 37; Indels 1; Gaps 1; | CC | The invention relates to a novel predictor set comprising a plurality of polynucleotides and/or polypeptides whose expression pattern is predictive of the response of cells to treatment with a compound that modulates protein tyrosine kinase activity or members of the protein tyrosine kinase pathway. The molecules of the invention demonstrate cytostatic, antiangiogenic, vasoconstrictive and vasoconstrictive activities and may be useful in the field of pharmacogenomics, in particular for determining drug sensitivity and in treating breast cancer, hypervascular diseases, angiogenesis and scars in wound healing. The current sequence is that of a human protein tyrosine kinase biomarker protein of the invention. |
| Qy | 1 MWLRVSSATVALLAQ-TELLFLVSRGPSSPAGGEARHVYLVLSSWRSSESSFVSQL 60 | CC | |
| Db | 1 MWLRVSSATVALLAQ-TELLFLVSRGPSSPAGGEARHVYLVLSSWRSSESSFVSQL 59 | CC | |
| Qy | 61 FSOHPDVYFLMEPAWHWNTLTSQGSAATLHMAVRLMRSIFLCIDMDVFDAYMPQSRNLSA 120 | CC | |
| Db | 60 FNQHPDVYFLMEPAWHWNTLTSQGSAATLHMAVRLMRSIFLCIDMDVFDAYLPWRNLSD 119 | CC | |
| SQ | Sequence 395 AA; | XX | |
| Qy | 121 FENWATSRALCSPPPACSAFPRTGSKQDVCKTLCTTROPSLAREACRSHVVLKEYRFF 180 | Query Match 84.3%; Score 1729.5; DB 8; Length 395; Best Local Similarity 85.8%; Pred. No. 2.4e-180; Matches 333; Conservative 17; Mismatches 37; Indels 1; Gaps 1; | |
| Db | 120 LFQMAVSRALCSPPPACSAFPRTGSKQDVCKTLCTTROPSLAREACRSHVVLKEYRFF 179 | CC | |
| Qy | 181 NLQVLYPLISDPALNRLIYHLVRDPRAVLSREOTAKALRDNGIVLGTNTGNKHVEADPHL 240 | Qy | 1 MWLPRESSSTVYLAWHHTTLSQGSAATHMAVRLMRSIFLCIDMDVFDAYMPOSRLSA 120 |
| Db | 180 NLQVLYPLISDPALNRLIYHLVRDPRAVLSREOTAKALRDNGIVLGTNTGNKHVEADPHL 239 | Db | 1 MWDPRSSVATALLAQ-TELLFLVSLRGPSPAGEGDRLVHVLVLSWRSGSSFVGQL 59 |
| Qy | 241 RLTVRECRSHVRTAEAATLKKPPPLGRGTRYLVPEDLAREPLAEIRALYAFTGLTTLTQL 300 | Qy | 61 FSQHPDVYFLMEPAWHWNTLTSQGSAATHMAVRLMRSIFLCIDMDVFDAYLPWRNLSD 119 |
| Db | 240 RVRECRSHVRTAEAATLKKPPPLGRGTRYLVPEDLAREPLAEIRALYAFTGLTTLTQL 299 | Db | 60 FNQHPDVYFLMEPAWHWNTLTSQGSAATHMAVRLMRSIFLCIDMDVFDAYLPWRNLSD 119 |
| Qy | 301 EAWHNITNGIGKPIEARHTSSRNARNVNSQMRHALPFTKLRLQEVCAAGALQLGYR 360 | Qy | 121 FENWATSRALCSPPPACSAFPRTGSKQDVCKTLCTTROPSLAREACRSHVVLKEYRFF 180 |
| Db | 300 EAWHNITNGIGKPIEARHTSSRNARNVNSQMRHALPFTKLRLQEVCAAGALQLGYR 359 | Qy | 122 FENWATSRALCSPPPACSAFPRTGSKQDVCKTLCTTROPSLAREACRSHVVLKEYRFF 179 |
| Qy | 361 PYYSDQQRDTLTDLVLPKGDFHFSWAS 388 | Db | 120 LFQMAVSRALCSPPPACSAFPRTGSKQDVCKTLCTTROPSLAREACRSHVVLKEYRFF 179 |
| Db | 360 PYYSDQQRDTLTDLVLPKGDFHFSWAS 387 | Qy | 181 NLQVLYPLISDPALNRLIYHLVRDPRAVLSREOTAKALRDNGIVLGTNGKHYEADPHL 240 |
| Qy | 180 NLQVLYPLISDPALNRLIYHLVRDPRAVLSREOTAKALRDNGIVLGTNGKHYEADPHL 239 | Db | 180 NLQVLYPLISDPALNRLIYHLVRDPRAVLSREOTAKALRDNGIVLGTNGKHYEADPHL 239 |
| Qy | 241 RLKEVCSHVRTAEAATLKKPPPLGRGTRYLVPEDLAREPLAEIRALYAFTGLTTLTQL 300 | Qy | 241 RLKEVCSHVRTAEAATLKKPPPLGRGTRYLVPEDLAREPLAEIRALYAFTGLTTLTQL 300 |
| Db | 240 RVRECRSHVRTAEAATLKKPPPLGRGTRYLVPEDLAREPLAEIRALYAFTGLTTLTQL 299 | Db | 240 RVRECRSHVRTAEAATLKKPPPLGRGTRYLVPEDLAREPLAEIRALYAFTGLTTLTQL 299 |
| Qy | 301 EAWHNITNGIGKPIEARHTSSRNARNVNSQMRHALPFTKLRLQEVCAAGALQLGYR 360 | Qy | 301 EAWHNITNGIGKPIEARHTSSRNARNVNSQMRHALPFTKLRLQEVCAAGALQLGYR 360 |
| Db | 300 EAWHNITNGIGKPIEARHTSSRNARNVNSQMRHALPFTKLRLQEVCAAGALQLGYR 359 | Db | 300 EAWHNITNGIGKPIEARHTSSRNARNVNSQMRHALPFTKLRLQEVCAAGALQLGYR 359 |
| RESULT 10 | | | |
| ID | ADL61235 | ID | AAV72638 standard, protein; 395 AA. |
| XX | standard; protein; 395 AA. | XX | AAV72638 |
| AC | ADL61235; | AC | AAV72638; |
| XX | | XX | |
| DT | 03-JUN-2004 (first entry) | DT | 02-MAY-2001 (first entry) |
| XX | | XX | |
| DE | Human tyrosine kinase biomarker carbohydrate sulphotransferase 6 protein. | DE | Mouse glycosyl sulfotransferase-4 (GST-4). |
| XX | predictor set; protein tyrosine kinase; cytostatic; antiangiogenic; | XX | Mouse; glycosyl sulfotransferase-4; GST-4; immunosuppressive; therapy; |
| KW | vasoconstrictive; vasoconstrictive; pharmacogenomic; drug sensitivity; breast cancer; | KW | selectin binding inhibitor; gene therapy; inflammation; |
| KW | hypervascular disease; angiogenesis; wound healing scar; human; | KW | systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes; |
| KW | biomarker; carbohydrate sulphotransferase 6; enzyme. | KW | polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis; |
| XX | Homo sapiens. | XX | glomerulonephritis; myasthenia gravis; Sjogren's syndrome; Hashimoto's disease; Graves' disease; hypoparathyroidism; anaemia; denervating disease; cirrhosis; ulcerative colitis; allergic rhinitis; myocarditis; duodenitis; respiratory distress syndrome; eczema; psoriasis; asthma; hypersensitivity; rheumatic fever; tissue rejection; |
| OS | | XX | chromosome 8E1. |
| XX | | XX | Mus musculus. |
| PN | WO2004020583-A2. | PN | |
| XX | 11-MAR-2004. | XX | |
| PD | | PD | |
| XX | 26-AUG-2003; 2003KC-US026491. | XX | |
| PP | | PP | |
| XX | 27-AUG-2002; 2002US-0406385P. | XX | |
| PR | | PR | |
| XX | (BRIM) BRISTOL-MYERS SQUIBB CO. | XX | |
| PA | | PA | |
| XX | Huang F, Han X, Reeves KA, Amler L, Fairchild CR, Lee FY; | XX | |
| PI | Shaw P; | PI | |
| XX | WPI; 2004-239171/22. | XX | |
| DR | | DR | |
| DR | N-PDB; ADL61098. | DR | |
| PT | New predictor sets with a plurality of polynucleotides and/or polypeptides whose expression pattern predicts cell response to a compound that modulates protein tyrosine kinase activity, useful in | PT | |
| PT | | PT | |

| | | |
|----------------------------------|---|---|
| PN | WO200106015-A1. | RESULT 11 |
| XX | PD 25-JAN-2001. | AAU11275 ID AAU11275 standard; protein; 395 AA. |
| XX | PP 19-JUL-2000; 2000WO-US019741. | XX |
| XX | PR 20-JUL-1999; 99US-0144694P. | AC AAU11275; |
| PR 13-JUN-2000; 2000US-00593828. | DR 12-MAR-2002 (first entry) | XX |
| XX | PA (REGC) UNIV CALIFORNIA. | DE Murine intestinal-GlcNAc 6-sulfotransferase (1-GlcNAc6S) protein. |
| XX | PI Rosen SD, Lee JK, Hemmerich S; | XX |
| XX | DR WPI; 2001-138471/14. | XX |
| XX | DR N-PSDB; ADD2696. | XX |
| XX | PT New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications. | XX |
| XX | PT diagnostic and therapeutic agent screening applications. | OS Mus musculus. |
| PS | Claim 3 ; Fig 2; 128pp; English. | XX |
| XX | CC The present sequence is mouse glycosyl sulfotransferase-4 (GST-4). GST-4 gene is found on chromosome 8E1. GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non-sulphated selectin ligand, GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's disease, Grave's disease, adrenalinitis, hypoparathyroidism, pernicious anaemia, demyelinating diseases, cirrhosis, ulcerative colitis, dermatitis, myocarditis, regional enteritis, adult respiratory distress syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation. | XX |
| XX | CC | XX |
| SQ | Sequence 395 AA; | PS 74.9%; Score 1536.5; DB 4; Length 395; |
| | Query Match Best Local Similarity 76.0%; Pred. No. 3.6e-159; Matches 295; Conservative 30; Mismatches 62; Indels 1; Gaps 1; | XX |
| Qy | 1 MWLPPRESSKTVILLAOTTCLLFLIFLRGPGSPAGGEDRVHVLSSWRSGSSFVQQL 60 | CC The present invention provides a method of modifying an acceptor molecule by contacting the acceptor with an isolated beta1,-N-acetylglucosaminyltransferase (beta1,3Gnt) or an active fragment, where beta1,3Gnt directs expression of a MECA-79 antigen. The invention also provides a method of treating or preventing an L-selectin-mediated condition by reducing the expression or activity of a beta1,3Gnt that directs expression of a MECA-79 antigen. This can be done by administering to the subject an oligosaccharide L-selectin antagonist that inhibits binding of L-selectin to a MECA-79 antigen, for example by administering antibody material that specifically binds beta1,3Gnt, and/or a beta1,3Gnt antisense nucleic acid molecule. L-selectin sulfotransferase-2 (USST-2), also directs MECA-79 antigen expression. Alternatively, the expression or activity of USST-2 or its active fragment can be reduced in combination with reducing the expression or activity of beta1,3Gnt. The method is useful for creating L-selectin mediated conditions such as Crohn's disease and ulcerative colitis, inflammatory disorders of the skin such as allergic contact dermatitis, psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type hypersensitivity reactions, diabetes and hyperplastic thymus. This sequence represents mouse I-GlcNAc6S |
| Db | 1 MRLPRESSTVMLSLMVGTVLFLVFLVRQPSSPAGLGERVHLVLSWRSGSSFVQQL 59 | XX |
| Qy | 61 FSOHPDVFLMEPAWHWVTLTSQGSAATLMAVRDLMSIIFLCDMDVFDAYMPQSRNLUSA 120 | XX |
| Db | 60 FSGHPDVFLMEPAWHWVNDLTSQGSAAPAHMAYVRDLIRSVLFCDMDVFDAYLPWRNISD 119 | CC |
| Qy | 121 FFWMTSRLCSPPACSFPPGTISKODVCKTLCTROPSLAREACRSYSHVLLKEVRF 180 | CC |
| Db | 120 LFQWAVSRALCSPPVCEAARGNISSEEVCPKLCATRPGLAQEAQCSYSHVLLKEVRF 179 | CC |
| Qy | 181 NLQVLYPLISDPAINLRLTVLHYLDPRAVLRSREBAGPLIARDNGIVLTGNGKWEADPHL 240 | CC |
| Db | 180 NLQVLYPLISDPAINLRLTVLHYLDPRAVLRSREQTAKLARDNGIVLTGNTWVEADPL 239 | CC |
| Qy | 241 RLREVCRSHVRLAEATAIKPPPLGRYLVRFEDLAREPLAEIRALYAFGLTLPQL 300 | CC |
| Db | 240 RVVNEVCVSHVRLAEAAHLKKPPLQDRYRLRYEDLARDPLTVIREYAFGLTLPQL 299 | CC |
| Qy | 301 EAWIHNITHGSIGKPKIAFHHTSSRNARNVSQAWRHLPFTKLRLVOEVCAAGLQLIGYR 360 | SQ Sequence 395 AA; |
| Db | 300 QTWIHNITHGGPQARREFKITSRDALSVSQAWRHTDFAKTRVQBLGGQLIGYR 359 | Query Match Best Local Similarity 76.0%; Pred. No. 3.6e-159; Matches 295; Conservative 30; Mismatches 62; Indels 1; Gaps 1; |
| Qy | 361 PYSADQORDLTIDLVLRGPDPFWSAS 388 | 1 MRLPRESSTVMLSLMVGTVLFLVFLVRQPSSPAGLGERVHLVLSWRSGSSFVQQL 60 |
| Db | 360 SVSBELEQRLSDLQLPRGMSDFKWA 387 | 1 MRLPRESSTVMLSLMVGTVLFLVFLVRQPSSPAGLGERVHLVLSWRSGSSFVQQL 59 |

CC which is given in comparison with (I) in the exemplification of the present invention

XX

QQ Sequence 418 AA;

DB Query Match 74.3%; Score 1523; DB 5; Length 418;

DB Best Local Similarity 75.8%; Pct. No. 1.2e-157;

DB Matches 294; Conservative 30; Mismatches 62; Indels 2; Gaps 2;

DB 1 MWLPRFESSRTVTLLQTCCLLFLISIIRGPESPAGGEEDDRVFLVLSSWRSGSSFVGQL 60

DB 25 MRDPRFESSVMTMLLMVGT-I-LVFLSRQVPSPAGSERVFLVLSSWRSGSSFVGQL 83

DB 61 FSQHPDVFYLMEPAWHTTILSGSAAATHMAYRDLMRSIIFLCMDMVDAYMPOSRNLSA 120

DB 84 FSQHPDVFYLMEPAWHTDTLSQGSAPALHNAVTRDLTRSVFLCDMDVDAYPARNLSD 143

DB 121 FENWATSRALCSPPPACSAFPRTGSKDYCKTLCTROPEFLAREACRSYSHVLFKEVRFF 180

DB 120 LFQWAVSRALCSPPPCEAFAARGNISSEEVKPLCATRPGLAQFCASSYSHVLFKEVRFF 179

DB 181 NLQVLYPLISDPALNRLTVHLVYDRAVLRSRREAGPILARDNGIVLGTNGKVAEADPHL 240

DB 180 NLQVLYPLISDPALNRLTVHLVDRAVLRSREQTAKALARDNGIVLGTNGKVAEADPHL 239

DB 241 RLINEVCRSHVRLIAEATLKKPPLRLGRYRLVREPDLLAREPLABIRALYAFTGTLTQL 300

DB 240 RVVNEVCRSHVRLIAEALHKPPPLQDRYTRLVYEDLARDPLTVIRELYAFTGIGLTQL 299

DB 301 EAWIHNIITGGIGKPKIAFHTISSRNARNVNSQAWRHLPFKILRVQEYVCAGAQQLQYR 360

DB 300 QTWIHNITGGSGPAGRREAKFTTSRDAVSQSQRHTLPFAKIRRVQEYLGGLQQLQYR 359

DB 361 PYVSDQDQLDTLVLPGPDHFSWAS 388

DB 360 SVHSELEQRDSLDDLPQMDSFKWAS 387

RESULT 12

ID ABB81557 standard; protein: 418 AA.

XX AC ABB81557;

XX DT 05-SEP-2002 (first entry)

XX DE Mouse intestinal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:5.

XX KW Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;

XX KW corneal; sulfatase; keratan sulfate; macular corneal dystrophy; MCD;

XX KW ophthalmological...

XX OS Mus musculus.

XX PN US2002061562-A1.

XX PD 23-MAY-2002.

XX PR 09-AUG-2001; 2001US-00927602.

XX PR 11-AUG-2000; 2000US-006318211.

XX PR 11-AUG-2000; 2000US-0325773P.

XX PA (FUKU/) FUKUDA M. N.

XX PA (AKAM/) AKAMA T. O.

XX PI Fukuda MN, Akama TO;

XX DR WPI: 2002-507643/54.

XX New nucleic acid encoding corneal N-acetylglucosamine-6-sulfotransferase, useful for treatment, monitoring and diagnosis of macular corneal dystrophy.

XX PS Example 5: Page 24-25; 69pp; English.

XX The present invention describes human corneal N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse sulfation of keratan sulfate (KS). Also described is a method for monitoring the effect of treatments for macular corneal dystrophy (MCD), and detecting susceptibility to MCD. (I) is located to chromosome 16q22, and has ophthalmological activity. (I) can be used to treat or prevent macular corneal dystrophy types I or II. (I) makes possible treatment of MCD without requiring keratoplasty or keratectomy. The present sequence represents mouse intestinal N-acetylglucosamine-6-sulfotransferase,

CC which is given in comparison with (I) in the exemplification of the present invention

XX

QQ Sequence 418 AA;

DB Query Match 74.3%; Score 1523; DB 5; Length 418;

DB Best Local Similarity 75.8%; Pct. No. 1.2e-157;

DB Matches 294; Conservative 30; Mismatches 62; Indels 2; Gaps 2;

DB 1 MWLPRFESSRTVTLLQTCCLLFLISIIRGPESPAGGEEDDRVFLVLSSWRSGSSFVGQL 60

DB 25 MRDPRFESSVMTMLLMVGT-I-LVFLSRQVPSPAGSERVFLVLSSWRSGSSFVGQL 83

DB 61 FSQHPDVFYLMEPAWHTTILSGSAAATHMAYRDLMRSIIFLCMDMVDAYMPOSRNLSA 120

DB 84 FSQHPDVFYLMEPAWHTDTLSQGSAPALHNAVTRDLTRSVFLCDMDVDAYPARNLSD 143

DB 121 FFWWATSRALCSPPPACSAFPRTGSKDYCKTLCTROPEFLAREACRSYSHVLFKEVRFF 180

DB 122 FFWWATSRALCSPPPACSAFPRTGSKDYCKTLCTROPEFLAREACRSYSHVLFKEVRFF 180

DB 144 LFQWAVSRALCSPPPVCEAFAARGNISSEBVCKPLCATRPGLAQEACSSYSHVLFKEVRFF 203

DB 181 NLQVLYPLISDPALNRLTVHLVYDRAVLRSRREAGPILARDNGIVLGTNGKVAEADPHL 240

DB 204 NLQVLYPLISDPALNRLTVHLVDRAVLRSREQTAKALARDNGIVLGTNGKVAEADPHL 263

DB 241 RLIREVCRSHVRLIAEATLKPPPLRLGRYRLVREPDLLAREPLABIRALYAFTGTLTQL 300

DB 264 RVVNEVCRSHVRLIAEALHKPPPLQDRYTRLVYEDLARDPLTVIRELYAFTGIGLTQL 322

DB 301 EAWHNITHGGIGKPKIAFHTISSRNARNVNSQAWRHLPFKILRVQEYVCAGAQQLQYR 360

DB 323 QTWIHNITHGGSGPAGRREAKFTTSRDAVSQSQRHTLPFAKIRRVQEYLGGLQQLQYR 382

RESULT 13

ID AD70405 standard; protein: 394 AA.

XX AC AD70405;

XX DT 06-MAY-2004 (first entry)

XX DE Human heart mitochondrial protein as a therapeutic target SeqID2211.

XX KW mitochrondrial; human; screening assay; diabetes mellitus;

XX KW Huntington's disease; osteoarthritis; Leber's hereditary optic neuropathy; LHON;

XX KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

XX KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

XX KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;

XX KW osteopathic; ophthalmologic; cytotatic.

XX OS Homo sapiens.

XX PN WO2003087768-A2.

XX PD 23-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010870.

XX PR 12-APR-2002; 2002US-0372843P.

XX PR 17-JUN-2002; 2002US-0369987P.

XX PR 20-SEP-2002; 2002US-0412418P.

XX PA (MITO-) MITOKOR.

XX PA (BUCK-) BUCK INST AGE RES.

XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

XX PI Warnock DE;

XX WPI; 2003-845369/78.
 DR PD 05-SEP-2001.
 XX PP 07-JUL-2000; 2000EP-00114089.
 PT XX
 PT 08-JUL-1999; 99JP-00194386.
 PT PR 11-JAN-2000; 2000JP-0011874.
 PT PR 02-MAY-2000; 2000JP-00183765.
 XX XX
 PT PA (HELI-) HELIX RES INST.
 XX XX
 PT Ota T, Nishikawa T, Isodai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuuki T, Koga H;
 XX XX
 PR WPI; 0011-524255/58.
 DR N-PSDB; AAF94229.
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PR in genetic manipulation.
 XX XX
 PR 830 Primers useful for synthesizing full length cDNA clones and their use
 PR in genetic manipulation.
 XX XX
 PR Claim 8; SEQ ID NO 2817; 1380pp + Sequence Listing; English.
 XX XX
 CC The invention relates to primers for synthesising full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesising the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a polypeptide encoded by a full length
 CC human cDNA of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in CD-ROM
 XX format directly from EPO.

Sequence 394 AA:
 Query Match 73.8%; Score 1513.5; DB 7; Length 394;
 Best Local Similarity 92.6%; Pred. No. 1.2e-156;
 Matches 289; Conservative 7; Mismatches 15; Indels 1; Gaps 1;
 Qy 1 MWLPRSSKTVTLLAQTTCUCLFLISRGPSSPAGGERVHLVLSMWSGSSFLGQL 60
 Db 1 MWLPRSSVATAILAQ-TFLLFLVLRGPSSPAGGERVHLVLSMWSGSSFLGQL 59
 SQ Sequence 386 AA:
 Query Match 73.8%; Score 1513.5; DB 7; Length 394;
 Best Local Similarity 92.6%; Pred. No. 1.2e-156;
 Matches 289; Conservative 7; Mismatches 15; Indels 1; Gaps 1;
 Qy 61 FSGQHPDVVEYLMEPAWHVVTTLSQCSAATLHMATDLMRSIELCDMDVYDAMYPOSRLNSA 120
 Db 60 FNQHPDVVEYLMEPAWHVVTTLSQCSAATLHMATDLMRSIELCDMDVYDAMYPOSRLNSD 119
 Qy 121 FFWVATSRALCSPPACSAPPRGFTSKQDVCKTLCTROPSLAREACSYSHVVLKEVRF 180
 Db 120 LFQAVAVSRALCSPPACSAPPRGFTSKQDVCKTLCARQPSSLAREACSYSHVVLKEVRF 179
 Qy 181 NLQVLYPLSDPAUNLRLVHLYRDPRAVLRSREAAQPLIARDNGIVIGTNGKWEADPHL 240
 Db 180 NLQVLYPLSDPAUNLRLVHLYRDPRAVLRSREAAQPLIARDNGIVIGTNGKWEADPHL 239
 Qy 241 RLIREVCRSHVRAAEAATLKKPPLGRYRLVRPEDLAREPLAEIRALYAFGLTLPQL 300
 Db 240 RLIREVCRSHVRAAEAATLKKPPLGRYRLVRPEDLAREPLAEIRALYAFGLTLPQL 299
 Qy 301 EAWHNITHGSG 312
 Db 300 EAWHNITHGSG 311
 RESULT 14
 ID AAM93309 standard; protein; 386 AA.
 XX AC AAM93309;
 XX AC
 XX DT 06-NOV-2001 (first entry)
 XX DE Human polypeptide, SEQ ID NO: 2817.
 XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 OS Homo sapiens.
 XX PN EP1130094-A2.
 XX XX
 RESULT 15
 ID ADL30784 standard; protein; 386 AA.
 ID ADL30784

| | | | | | |
|---------|---|---|---------------------------------|---|-------------------|
| AC | ADL30784 | 299 | LPHLQTVVHNITRGKGMD- | HAFHTNARDALNVSQAWRWSLPEKYVSRLQKACGDAANL | 356 |
| XX | DT 20-MAY-2004 (first entry) | 357 | LGYRPVYSAQQDLDLDDVLPGPDHFSWASPD | 390 | |
| XX | Human protein encoded by a full length cDNA clone SedID 2817. | 357 | LGYRHRVSEQEQRNLLDL-----STWTVPE | 383 | |
| DE | | | | | |
| XX | human; medicine; signal transduction; glycoprotein; transcription; | | | | |
| KW | oligo-capping method. | | | | |
| XX | | | | | |
| OS | Homo sapiens. | | | | |
| XX | | | | | |
| PN | EP1396543-A2. | | | | |
| XX | | | | | |
| PD | 10-MAR-2004. | | | | |
| XX | | | | | |
| PP | 07-JUL-2000; 2003EP-00025638 | | | | |
| XX | | | | | |
| PR | 08-JUL-1999; 99JP-00194486 | | | | |
| PR | 11-JAN-2000; 2000JP-00118774 | | | | |
| PR | 02-MAY-2000; 2000JP-00183865 | | | | |
| PR | 07-JUL-2000; 2000EP-00114089 | | | | |
| XX | | | | | |
| PA | (REAS-) RES ASSOC BIOTECHNOLOGY. | | | | |
| XX | | | | | |
| PI | Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y; | | | | |
| PI | Wakamatbu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H; | | | | |
| XX | | | | | |
| DR | WPI; 2004-204755/20. | | | | |
| DR | N-PSDB; ADL30783. | | | | |
| XX | | | | | |
| PT | New oligonucleotide primers (830 cDNAs) useful for synthesizing full | | | | |
| PT | length human cDNAs. | | | | |
| XX | | | | | |
| PS | Example 1; SEQ ID NO 2817; 1340pp; English. | | | | |
| XX | | | | | |
| CC | This invention relates to a novel primers useful for synthesising full | | | | |
| CC | length cDNA molecules that encode human proteins. Specifically, it refers | | | | |
| CC | to secretory or membrane proteins that are potential therapeutic agents/ | | | | |
| CC | target molecules in the field of medicine, and in particular genes | | | | |
| CC | encoding proteins that are associated with signal transduction, | | | | |
| CC | glycoproteins and transcription. The present invention describes a method | | | | |
| CC | for efficiently cloning a full length human cDNA from both the 5' and 3' | | | | |
| CC | ends using the oligo-capping method. This polypeptide sequence is a full | | | | |
| CC | length human protein of the invention. | | | | |
| XX | | | | | |
| SQ | Sequence 386 AA; | | | | |
| | | | | | |
| Query | Match | 49.4% | Score | 1013; | DB 8; Length 386; |
| Best | Local Similarity | 52.5% | Pred. | No. 9.9e-102; | |
| Matches | Conservative | 55; | Mismatches | 104; | Indels 7; |
| | | | | | |
| Qy | 14 | LLAQTTCLLIFISRP----- | GPSSPAGGEDRIVHVLSSWRSSSSFVQ | 59 | |
| Db | 1 | MLLPKMKMFLVQMAILAFFFHMYSNISSLSMAQERHMYVLVLSWRSSSFVQ | 60 | | |
| Qy | 60 | LFSQHPDVFLMPEAMHWTTLQSQSATTLMARDLMSIPLCDMDVSDAYM PQSRNL | 118 | | |
| Db | 61 | LFQHPDVFLMPEAWHMTEFKQSTAMHMAVDLIRAVFLCDMSVTFAYM PGPRIQ | 120 | | |
| Qy | 119 | SAFFWATSRALCSPPACSFPRGTISKQDYCKTLCTQPSLAREACSYSHVVKERYR | 178 | | |
| Db | 121 | SSLFOWENSRALCSAPACDIPQDDELIPIRAHCRLCSQQPFEVVKACRSYSHVVKERYR | 180 | | |
| Qy | 179 | FFNLOQLYPISDPAINLRYHVLDPRAVIRSREAAQPLARDNGIVGTN-CKWVEAD | 237 | | |
| Db | 181 | FFNLOQLYPILKDPSPNLHVTYHVLDPRAVFRSRTKGDIMIDRIVNCQHEQLKED | 240 | | |
| Qy | 238 | PHRLTRECVCSHVRIAETATLKP-PPFLKGRYLVRFEDLAIREALYAFGTTL | 296 | | |
| Db | 241 | QPYVMQVICOSQLEYK-TIQSLSPKALDQERYLLVRYEELARAQAQTSMYEFVGLEF | 298 | | |
| Qy | 297 | TPOLEAWHNTTHGSIGKPKTEAFTHTSSRNARNVSOAHRHALPFTKILRVQEVCAGALQ | 356 | | |
| Db | | | | | |

THIS PAGE BLANK (USPS TO)

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 08:33:13 ; Search time 42.7985 Seconds
 (without alignments)
 46666.302 Million cell updates/sec

Title: US-10-697-828-8

Perfect score: 2051

Sequence: 1 MWLPRFSSKTVTLLAQTTF.....LTLIDVLVLRGPDHFSWASPD 390

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03:
 1: uniprot_sprot:
 2: uniprot_trembl:
 *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB ID | Description |
|------------|-------|-------|-------|--------|------------|-----------------------------|
| 1 | 2051 | 100.0 | 390 | 2 | Q7LCN3 | Q7lcn3 homo sapien |
| 2 | 2051 | 100.0 | 411 | 2 | Q9GZS9 | Q9gzs9 homo sapien |
| 3 | 1729 | 5.5 | 84.3 | 395 | Q9GZX3 | Q9gzx3 homo sapien |
| 4 | 1536 | 5.5 | 74.9 | 395 | Q9QUP4 | Q9qup4 mus musculus |
| 5 | 1119 | 5.5 | 54.6 | 392 | Q6GN39 | Q6gn39 xenopus lae |
| 6 | 1013 | 4.9 | 49.4 | 386 | Q8NCG5 | Q8ncg5 homo sapien |
| 7 | 1008 | 4.9 | 49.1 | 386 | Q9Y5R3 | Q9y5r3 homo sapien |
| 8 | 999 | 48.7 | 370 | 2 | Q8IV46 | Q8iv46 homo sapien |
| 9 | 965 | 47.1 | 388 | 2 | Q9R1I1 | Q9r1i1 mus musculus |
| 10 | 958 | 46.7 | 388 | 2 | Q9WU5 | Q9wu5 mus musculus |
| 11 | 665 | 32.4 | 483 | 2 | Q994G9 | Q994g9 mus musculus |
| 12 | 665 | 32.4 | 530 | 2 | Q8B2T6 | Q8b2t6 mus musculus |
| 13 | 662 | 32.3 | 530 | 2 | Q80WV3 | Q80wv3 mus musculus |
| 14 | 659 | 32.2 | 483 | 2 | Q9UED3 | Q9ued3 homo sapien |
| 15 | 659 | 32.2 | 530 | 2 | Q9Y4C5 | Q9y4c5 homo sapien |
| 16 | 617 | 30.1 | 484 | 2 | Q9EP78 | Q9ep78 m-n-acetylgl |
| 17 | 614 | 30.0 | 484 | 2 | Q9NB0 | Q9nb0 mus musculus |
| 18 | 610 | 29.7 | 486 | 2 | Q75667 | Q75667 homo sapien |
| 19 | 609 | 29.7 | 486 | 2 | Q9NS84 | Q9ns84 homo sapien |
| 20 | 603 | 29.4 | 485 | 2 | Q6XQGB | Q6xqgb rattus norvegicus |
| 21 | 579 | 28.3 | 479 | 2 | Q7LGCB | Q7lgcb homo sapien |
| 22 | 577 | 28.2 | 479 | 2 | Q75099 | Q75099 homo sapien |
| 23 | 576 | 28.1 | 472 | 2 | Q8199 | Q8199 mus musculus |
| 24 | 560 | 27.3 | 474 | 2 | Q9QZL2 | Q9qzl2 rattus norvegicus |
| 25 | 552 | 26.9 | 420 | 2 | Q6DBY9 | Q6dby9 brachydanio |
| 26 | 540 | 26.5 | 485 | 1 | Q6ST_CHICK | Q6st_chick gallus gallus |
| 27 | 520 | 25.4 | 411 | 2 | Q9E0C0 | Q9e0c0 mus musculus |
| 28 | 513 | 25.0 | 411 | 2 | Q43916 | Q43916 homo sapien |
| 29 | 472 | 23.0 | 441 | 2 | Q79403 | Q79403 torpedo californicus |
| 30 | 403 | 19.7 | 257 | 2 | Q79415 | Q79415 mus musculus |
| 31 | 353 | 17.2 | 225 | 2 | Q6RY62 | Q6ry62 cavia porcellus |

ALIGNMENTS

| RESULT 1 | | | | | | | | | |
|----------|---------|--|--|--|--|--|--|--|--|
| Q7LCN3 | | | | | | | | | |
| ID | Q7LCN3 | PRELIMINARY; | | PRT; | | 390 AA. | | | |
| AC | Q7LCN3; | (TREMBLrel. 27, Created) | | DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update) | | DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update) | | | |
| DE | | N-acetylglucosamine-6-S- | | RT GN Name=I-GLCNAC-6-ST; | | RT OC Homo sapiens (Human) | | | |
| | | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butteleostomi; | | RT OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | RT NCBI_TaxID=9606; | | | |
| RN | | {[1]} | | RN SEQUENCE FROM N.A. | | RN SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RN | | {[2]} | | RN SEQUENCE FROM N.A. | | RN SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RNA | | {[3]} | | RNA SEQUENCE FROM N.A. | | RNA SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RNA | | {[4]} | | RNA SEQUENCE FROM N.A. | | RNA SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RNA | | {[5]} | | RNA SEQUENCE FROM N.A. | | RNA SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RNA | | {[6]} | | RNA SEQUENCE FROM N.A. | | RNA SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RNA | | {[7]} | | RNA SEQUENCE FROM N.A. | | RNA SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RNA | | {[8]} | | RNA SEQUENCE FROM N.A. | | RNA SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | | |
| RC | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | RC TISSUE=Intestine; | | | |
| RA | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RT | | Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | RT Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; | | | |
| RL | | "Cloning and characterization of a mammalian N-acetylglucosamine-6-Sulfotransferase that is highly restricted to intestinal tissue." | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | RL Biochem. Biophys. Res. Commun. 263:543-549 (1999). | | | |
| RNA | | {[9]} | | RNA SEQUENCE FROM N.A. | | RNA SEQUENCE FROM N.A. | | | |
| RP | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A. | | RP SEQUENCE FROM N.A.</ | | | |

| | | | |
|---|---|---|---|
| DR | EMBL; BC074983; AAH74983.1; -; | DR | MGD; MGI:1931825; Chats5. |
| DR | BC074984; AAH74984.1; -; | DR | GO; GO:0005615; C:extracellular space; TAS. |
| DR | AF219930; AAG26353.1; -; | DR | InterPro; IPR000663; Sulfotransferase. |
| DR | Gene; HGNC:6938; CHST6. | DR | PFam; PF00685; Sulfotransferase_1. |
| DR | GO; GO:001574; C:Golgi apparatus; TAS. | SQ | SEQUENCE 395 AA; 44537 MW; 3DFD71B43ED383BE CRC64; |
| DR | GO; GO:0001517; P:N-acetylglucosamine 6-O-sulfotransferase ac...; TAS. | SEQUENCE 395 AA; 44098 MW; 433CA60248A48F67 CRC64; | Query Match 74.9%; Score 1536.5; DB 2; Length 395; |
| DR | GO; GO:0006044; P:N-acetylglucosamine metabolism; IC. | Best Local Similarity 76.0%; Pred. No. 4; se-129; | Best Local Similarity 76.0%; Pred. No. 4; se-129; |
| DR | InterPro; IPR000863; Sulfotransferase. | Matches 295; Conservative 30; Mismatches 62; Indels 1; Gaps 1; | Matches 295; Conservative 30; Mismatches 62; Indels 1; Gaps 1; |
| DR | PFam; PF00685; Sulfotransferase_1. | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| KW | KW Transferease. | Db 1 MWLPRFSSTVMLSLMMWQTG-ILVFLVSRQVPPSSPAGLGERHVVLSSWRSGSSFVGQL 59 | Db 1 MWLPRFSSTVMLSLMMWQTG-ILVFLVSRQVPPSSPAGLGERHVVLSSWRSGSSFVGQL 59 |
| SQ | SEQUENCE 395 AA; 44098 MW; 433CA60248A48F67 CRC64; | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Query Match 84.3%; Score 1729.5; DB 2; Length 395; | Best Local Similarity 85.8%; Pred. No. 2.2e-146; Indels 1; Gaps 1; | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Matches 333; Conservative 17; Mismatches 37; Indels 1; | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 59 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Qy 61 FSOHPDVYLMEPAWHVVNTLSSQSAATLHMAVRDLMSIFLCIDMDVFDAYMPOSRLNSA 120 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 60 FNQHPDVYLMEPAWHVVNTLSSQSAATLHMAVRDLMSIFLCIDMDVFDAYLPPRNNSD 119 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Qy 121 FENWATSRALCSPPACSAFPRTGTSKODVCKTLCTROPSLAREACRSYSHVVKERYFF 180 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 120 LFQAVSRAALCSPPACSAFPRTGTSKODVCKTLCTROPSLAREACRSYSHVVKERYFF 180 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 120 LFQAVSRAALCSPPACSAFPRTGTSKODVCKTLCTROPSLAREACRSYSHVVKERYFF 179 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Qy 181 NLQVLYPLISDPALNLRIVLVRDPAVLRSREAAQPLCARQSLSEAVCPKLAQSLVYKERYF 240 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 180 NLQVLYPLISDPALNLRIVLVRDPAVLRSREAAQPLCARQSLSEAVCPKLAQSLVYKERYF 239 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Qy 241 RLIREVCRSVHRIAATLKKPPPLGRYRVLVPEDLARPLAFTIALYTAFTGLTLPQL 300 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 240 RVVIEVCRSVHRIAATLKKPPPLGRYRVLVPEDLARPLAFTIALYTAFTGLTLPQL 299 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Qy 301 EAIIHNITIGSGIGKPIEAFTTSSRNARYSQAMRHAPFTKILRQEYCAAGALQLGYR 360 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 300 EAIIHNITIGSGIGKPIEAFTTSSRNARYSQAMRHAPFTKILRQEYCAAGALQLGYR 359 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Qy 361 PVYSADQQDLTDLIVLPKGPDHFWSMAS 388 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Db 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| Db 360 PVYSSEDEQRNLALDILVLPKGNGFTWAS 387 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 | Qy 1 MWLPRFSSKTVVLLAQTCLLFIISRGPGPSSPAGEDRVRVTLVLSWRSGSSFLGQL 60 |
| RESULT 4 Q9QP4 | SEQUENCE FROM N.A. Q9QP4; PRELIMINARY; 395 AA. | Q6GN39 PRELIMINARY; PRT; 392 AA. | SEQUENCE FROM N.A. Q6GN39 PRELIMINARY; PRT; 392 AA. |
| ID Q9QP4 | AC Q9QP4; PRELIMINARY; 395 AA. | ID Q6GN39 ID Q6GN39; AC Q6GN39; GN Q6GN39; DT 05-JUL-2004 (TREMBrel. 27, Created) | SEQUENCE FROM N.A. Q6GN39 ID Q6GN39; AC Q6GN39; GN Q6GN39; DT 05-JUL-2004 (TREMBrel. 27, Last sequence update) |
| DT 01-MAY-2000 (TREMBrel. 13, Created) | DT 01-MAY-2000 (TREMBrel. 13, Last sequence update) | DT 05-JUL-2004 (TREMBrel. 27, Last annotation update) | DT 05-JUL-2004 (TREMBrel. 27, Last annotation update) |
| DT 05-JUL-2004 (TREMBrel. 27, Last annotation update) |
| DE Name=Chst5; Synonyms=I-GlcNAc-6-SN; | DE Name=Chst5; Synonyms=I-GlcNAc-6-SN; | DE Name=Chst5; Synonyms=I-GlcNAc-6-SN; | DE Name=Chst5; Synonyms=I-GlcNAc-6-SN; |
| OS Mus musculus (Mouse) |
| OC Mammalia; Eutheria; Rodentia; Chordata; Craniata; Vertebrata; Buteleostomi; Murinae; Mus. | OC Mammalia; Eutheria; Rodentia; Chordata; Craniata; Vertebrata; Buteleostomi; Murinae; Mus. | OC Mammalia; Eutheria; Rodentia; Chordata; Craniata; Vertebrata; Buteleostomi; Murinae; Mus. | OC Mammalia; Eutheria; Rodentia; Chordata; Craniata; Vertebrata; Buteleostomi; Murinae; Mus. |
| RN Q9QP4 | RN Q9QP4 | RN Q9QP4 | RN Q9QP4 |
| RN SEQUENCE FROM N.A. Q9QP4; PRELIMINARY; 395 AA. | RN SEQUENCE FROM N.A. Q9QP4; PRELIMINARY; 395 AA. | RN SEQUENCE FROM N.A. Q9QP4; PRELIMINARY; 395 AA. | RN SEQUENCE FROM N.A. Q9QP4; PRELIMINARY; 395 AA. |
| RC STRAIN=C7B1/6; TISSUE=Intestine; MEDLINE=9423499; PubMed=1041328; DOI=10.1006/bbrc.1999.1324; |
| RX Lee J.K., Bhakta, S., Rosen S.D., Hemmerich S.; "Cloning and characterisation of a mammalian N-acetylglucosamine-6-sulotransferase that is highly restricted to intestinal tissue."; Biochem. Biophys. Res. Commun. 263:543-549 (1999). | RX Lee J.K., Bhakta, S., Rosen S.D., Hemmerich S.; "Cloning and characterisation of a mammalian N-acetylglucosamine-6-sulotransferase that is highly restricted to intestinal tissue."; Biochem. Biophys. Res. Commun. 263:543-549 (1999). | RX Lee J.K., Bhakta, S., Rosen S.D., Hemmerich S.; "Cloning and characterisation of a mammalian N-acetylglucosamine-6-sulotransferase that is highly restricted to intestinal tissue."; Biochem. Biophys. Res. Commun. 263:543-549 (1999). | RX Lee J.K., Bhakta, S., Rosen S.D., Hemmerich S.; "Cloning and characterisation of a mammalian N-acetylglucosamine-6-sulotransferase that is highly restricted to intestinal tissue."; Biochem. Biophys. Res. Commun. 263:543-549 (1999). |
| RX 1 [2] | RX 1 [2] | RX 1 [2] | RX 1 [2] |
| RN STRAIN=C7B1/6; TISSUE=Intestine; MEDLINE=9423499; PubMed=1041328; DOI=10.1006/bbrc.1999.1324; |
| RA Lee J.-K., Heiton E., Kerteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko A., Blakesley R.W., Touchman J.W., Green B.D., Bouffard G.G., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., | RA Lee J.-K., Heiton E., Kerteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko A., Blakesley R.W., Touchman J.W., Green B.D., Bouffard G.G., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., | RA Lee J.-K., Heiton E., Kerteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko A., Blakesley R.W., Touchman J.W., Green B.D., Bouffard G.G., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., | RA Lee J.-K., Heiton E., Kerteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko A., Blakesley R.W., Touchman J.W., Green B.D., Bouffard G.G., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., |
| RA Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases. |
| DR EMBL; API76841; AA056003.1; |
| DR | DR | DR | DR |

RESULT 6

| | | | | |
|----|--|--------------|------|---------|
| ID | QBNCG5 | PRELIMINARY; | PRT; | 386 AA. |
| AC | QBNCG5; | | | |
| DT | 01-OCT-2002 (TrEMBLrel. 22, Created) | | | |
| DT | 01-OCT-2002 (TrEMBLrel. 22, Last sequence update) | | | |
| DE | Hypothetical protein FLJ90265. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TAXID=9606; | | | |
| RN | SEQUENCE FROM N.A. | | | |
| RA | Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.B., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [2] | | | |
| RC | SEQUENCE FROM N.A. | | | |
| RA | TISSUE=embryo; | | | |
| RA | Medline:22341132; PubMed:12454917; DOI:10.1002/dvdy.10174; | | | |
| RA | Klein S.L., Strauberg R.L., Wagner L., Pontius J., Clifton S.W., Richardson P.; "Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative."; Dev. Dyn. 225:384-391 (2002). [3] | | | |
| RA | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=embryo; | | | |
| RA | Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases. | | | |
| RL | DR GO: 0008146; F:sulfotransferase activity; IEA. | | | |
| DR | InterPro; IPR00063; Sulfotransferase. | | | |
| DR | PFAM; PF00685; Sulfotransferase_1. | | | |
| SQ | SEQUENCE 392 AA; 45969 MW; 45160 MW; 861869348319E42A CRC64; | | | |

Query Match 49.4%; Score 1013; DB 2; Length 386;

Best Local Similarity 52.5%; Pfd. No. 3.7e-92;

Matches 207; Conservative 55; Mismatches 104; Indels 28; Gaps 7;

Query 14 LLLAQTTCLLPLIISRP-----GPSSPAGGEDDRVHVLVSSWRSGSSFLGQ 59

Db 1 MLLPKKMLFLVSSWMAILALFEMYSHNTSSLMAOPPERMHVLSRSGSSFLGQ 60

Query 60 LFSQHPDVYLMEPANWVWTLQSASATLMAVRDLMRSFLCLDMDFDAYM-POSRNL 118

Db 61 LFGQHPDVYLMEPANWVWTLQKOSTPAWMHMAVRDLMIRAVFLCDMSVEDAYMEPGRRQ 120

Query Match 49.4%; Score 1013; DB 2; Length 386;

Best Local Similarity 52.5%; Pfd. No. 3.7e-92;

Matches 207; Conservative 55; Mismatches 104; Indels 28; Gaps 7;

Query 14 LLLAQTTCLLPLIISRP-----GPSSPAGGEDDRVHVLVSSWRSGSSFLGQ 59

Db 1 SLLFQWNSRALCSAPCDITPQDEITPACRLLSQQPFRVEXACRSFTSHVVLKEVR 180

Query 179 FPNLQLVYPLLSDPALNRLIVLVRDRAVRSREAAQPLIARDNGVILGN-GKWEAD 237

Db 181 FPNLQLVYPLLSDPALNRLIVLVRDRAVRSREAAQPLIARDNGVILGN-GKWEAD 240

Query 238 PHRLRLTRECRSHVRIAETATLKP-BPFLGRYLYVRFEDAREPLAEIRALYAFGTGLT 296

Db 241 QPYYVMQVICQSLEYIK-TIQSLSPXALQERYLLTVYEDLARAP/QTSRMVEFYQLEF 298

Query 297 TPOLEAWHNITHGSGIGKPKIAAHTSRRNARNVSQMRHALPFTKLRLVQEVCAQALQ 356

Db 299 LPHLQTWHNTRGKGKD-HAFTNARDALNVSQMRWLSPYEVSRLOKACGDMNL 356

Query 357 LGYRPVYSAQDQDLTLIDLVLRGPDPHFSWASPD 390

Db 357 LGYRHVSEQEBRNLLDIL-----STWTYPE 383

RESULT 7

| | | | |
|--------|--|------|---------|
| Q9Y5R3 | PRELIMINARY; | PRT; | 386 AA. |
| AC | O9Y5R3 | | |
| DT | 01-NOV-1999 (TrEMBLrel. 12, Created) | | |
| DT | 01-NOV-1999 (TrEMBLrel. 12, Last sequence update) | | |
| DT | 05-JUL-2004 (TrEMBLrel. 27, Last annotation update) | | |
| DE | N-acetylglucosamine 6-O-sulfotransferase (L-selectin ligand sulfotransferase GST-3). | | |
| GN | Name=SST3; | | |
| OS | Homo sapiens (Human). | | |
| OC | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| OX | NCBI_TAXID=9606. | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | TrEMBL; PMID=10330415; DOI=10.1083/jcb.145.4.899; | | |
| RA | Bistrip A., Bhakta S., Lee J.X., Belov Y.Y., Gunn M.D., Zuo F.R., Huang C.C., Kanngi R., Rosen S.D.; Hemmerich S.; "Sulfotransfases of two specificities function in the reconstitution of high endothelial cell ligands for L-selectin."; | | |
| RT | J. Cell Biol. 145:895-910 (1999). | | |
| [1] | | | |
| [2] | | | |
| RA | SEQUENCE FROM N.A. | | |
| RA | Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T., | | |
| RA | SEQUENCE FROM N.A. | | |
| RA | SEQUENCE FROM N.A. | | |

| | |
|---|---|
| GN Name=Chst4; | DB 271 LFLRYEDLVRPLAQTTRLYKFVGLDFLPHIQTWWNVTRGKGMQ--HAFHTNARNALN 328 |
| OS Mus musculus (Mouse). | QY 330 VSQAWRHALPFKILRVQEYCAAGALOLGYPRVYSSADQORDLTLDLV 376 |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | DB 329 VSQAWRNSLPYBKVSQLDACGEAMDILGYLQRSSOBQGNLSLDLL 375 |
| OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus. | |
| NCBI_TAXID=10090; | |
| RN [1] | |
| SEQUENCE FROM N.A. | RESULT 10 |
| RX STRAIN=PVB/N; TISSUE=Colon; | SEQUENCE FROM N.A. |
| RX PubMed=1477932; DOI=10.1073/pnas.242603899; | STRAIN=C57BL/6; TISSUE=Tongue; |
| RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., | MEDLINE=99264336; PubMed=0330415; DOI=10.1083/jcb.145.4.899; |
| RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schulter G.D., | RA Bistrup A., Bhakta S., Lee J.K., Belov Y.Y., Gunn M.D., Zuo F.R., |
| RA Aitschbaud S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., | RA Huang C.C., Kamagi R., Rosen S.D., Hemmerich S., |
| RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., | RT "Sulfotransfases of two specificities function in the reconstitution |
| RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.J., Abramson R.D., Mulahay S.J., | RT of high endothelial cell ligands for L-selectin."; |
| RA Raha S.S., Loqueilano N.A., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., | RL J. Cell Biol. 145:899-910(1999). |
| RA Richards S., McEwan P.J., McKernan K.J., Malek A.M., Gay L.J., Hulyk S.W., | RN [2] |
| RA Villalon D.K., Murzy D.M., Sodergren B.J., Lu X., Gibbs R.A., | RP SEQUENCE FROM N.A. |
| RA Fahy J., Helton E., Ketteman A.C., Rodriguez S., Sanchez A., | STRAIN=C57BL/6J; TISSUE=Tongue; |
| RA Whiting M., Madan A., Young A.C., Shevchenko A., Bouffard G.G., | MEDLINE=99279255; PubMed=0349636; DOI=10.1016/S0076-6879(99)03004-9; |
| RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | RA Carnici P., Hayashizaki Y., |
| RA Rodriguez A.C., Grimmwood J., Myers R.M., Butterfield Y.S., | RT High-efficiency full-length cDNA cloning."; |
| RA Kizuyiwaki M.I., Skalska U., Smaltz D.E., Schniech A., Schein J.B., | RL Meth. Enzymol. 303:19-44(1999). |
| RA Jones S.J., Marra M.A.; | RN [3] |
| RA "Generation and initial analysis of more than 15,000 full-length human | RP SEQUENCE FROM N.A. |
| RT and mouse cDNA sequences."; | STRAIN=C57BL/6J; TISSUE=Tongue; |
| RT Submitted (SPB-2003) to the EMBL/GenBank/DBJ databases. | MEDLINE=21085660; PubMed=11217851; DOI=10.1018/35055500; |
| DR EMBL: AF0109155; AAD45579.1; -. | RA RIKEN FANTOM Consortium; |
| DR EMBL: BC057886; AHH7886.1; -. | RT "Functional annotation of a full-length mouse cDNA collection."; |
| DR MGDN: MGII:1349479; Chst4. | RL Nature 409:685-690 (2001). |
| DR GO: GO:0016021; C:integral to membrane; TAS. | RN [4] |
| DR InterPro: IPR000863; Sulfotransfase. | RP SEQUENCE FROM N.A. |
| DR Pfam: PF00685; Sulfotransfer_1; 1. | STRAIN=C57BL/6J; TISSUE=Tongue; |
| KW Selectin; Transferase. | RA The RIKEN Genome Exploration Research Group Phase I & II Team; |
| SQ SEQUENCE FROM N.A. | RA "Analysis of the mouse transcriptome based on functional annotation of |
| Query Match 47.1%; Score 965.5; DB 2; | RT 60,770 full-length cDNAs."; |
| Best Local Similarity 55.9%; Pred. No. 6.0e-78; | RL Nature 420:563-573 (2002).; |
| Matches 194; Conservative 41; Mismatches 87; | RN [5] |
| Indels 25; Gaps 6; | RP SEQUENCE FROM N.A. |
| Score 44635 MW; 6D5371AFB6884ABE CRC64; | STRAIN=C57BL/6J; TISSUE=Tongue; |
| Query 42 VHVVLSSRSRGSGSFLGOLFQHPDVFLVNLQYLPSDPAHLRSLRSIF 101 | MEDLINE=2049937; PubMed=1042159; DOI=10.1101/gr.145100; |
| DB 42 VHVVLSSRSRGSGSFLGOLFQHPDVFLVNLQYLPSDPAHLRSLRSIF 101 | RA Shibusawa K., Akiyama J., Nisika K., Nagatoka S., Sasai N., Carninci P., |
| DB 102 LCDMDVFDAYM-POSRNISAFFWATSDALCSPPACSAPFRGTTSKODVCKTLCTROFS 160 | RA Konno H., Itoh M., Aizawa K., Harada A., Itoh M., |
| DB 102 LCDMSVFDAYMNPGPKQSSLFQWEQSRALCSAPVCDFFPAHEISSPKHCKLCGQQPFD 161 | RA Sumi N., Ishii Y., Nakamura S., Hazama M., Sugahara Y., Shibata K., Itoh M., |
| QY 161 LARFACRSTSYTSHVILKEVRFNQLQYLPSDPAHLRSLRSIF 101 | RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., |
| DB 162 MVEACRSGFVFLKEVRFNQLQYLPSDPAHLRSLRSIF 101 | RA Fujiwara S., Inoue K., Togawa Y., Ohara E., Watshiki M., |
| DB 221 RDNGIVLTGNGKWEADPHLIRE-----VCRSHVRIAEAA-TLKKPPFLGRY 269 | RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J., |
| QY 102 LCDMDVFDAYM-POSRNISAFFWATSDALCSPPACSAPFRGTTSKODVCKTLCTROFS 160 | |
| DB 102 LCDMSVFDAYMNPGPKQSSLFQWEQSRALCSAPVCDFFPAHEISSPKHCKLCGQQPFD 161 | |
| QY 161 LARFACRSTSYTSHVILKEVRFNQLQYLPSDPAHLRSLRSIF 101 | |
| DB 162 MVEACRSGFVFLKEVRFNQLQYLPSDPAHLRSLRSIF 101 | |
| DB 222 VDSHIVLG-----QHLETIKEEDQPYANKIICKSQDVTKAQTLL--PEALQORY 270 | |
| QY 270 RLVRPFDLAREPLAETRAYAFTGLTLQPOLEWIHNTHGSGIGKPIEAFFTSRNARN 329 | |

PA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RISA integrated sequence analysis (RISA) system-384-format"
 RT sequencing pipeline with 384 multicapillary sequencer.";
 Genome Res. 10:1157-1171(2000).

SEQUENCE FROM N.A.

RC SPBAIN-C57BL/6J; TISSUE=Tongue;

RA Adachi J., Akahira S., Akimura T., Arai A., Aono H., Furuno M.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Hori F.,
 RA Hanagaki T., Hara A., Hayashi N., Hiramoto K., Hirooka T., Hori M.,
 RA Imotani K., Ishii M., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Nunazaki R., Ohno M.,
 RA Okada Y., Okido T., Owa C., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibusawa Y., Shiraki T.,
 RA Segabe Y., Suzuki H., Tagami M., Takahashi F., Tanaka T.,
 RA Tejima T., Toya T., Yamamoto T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (TUL-2000) to the EMBL/GenBank/DDBJ databases.
 EMBL: AK091133; BAB26078.1; -.

DR MGDB; MGJ:1349479; Chst4.

DR GO:0016021; C:integral to membrane; TAS.

DR InterPro: IPR000833; Sulfotransferase_1.

DR PFam: PF00685; Sulfotransferase_1.

KW Transferase.

SEQUENCE 388 AA: 44694 MW: D1E9D7796DF8574D CRC64;

Query Match 46.7%; Score 958.5; DB 2; Length 388;
 Best Local Similarity 55.6%; Pred. No. 2.9e-77;
 Matches 193; Conservative 42; Mismatches 87; Indels 25; Gaps 6;

Qy 42 VHVLVLSMRSGSSLFLGOLFQSQHPDVFYLMEPAWHVVWTLSQGSAATHMAYRDLSIF 101
 Db 42 VHVVLSSMRSGSSPVGQLFGQHPDVFYLMEPAWHVVMTFTSSTAWKLMAYRDLLRSVF 101

Qy 102 LCMDVFDAYM-P-QSRNLSAFFENPATSRALCSPPACSAFPRTGTSKODVKTLCTROFPS 160
 Db 102 LCDNSVFDAYMNPGRKQSLFQBEOSRALSACAVCDFFPAHEBSSPKHCKLGQQFD 161

Qy 161 LAREACRSYSHVILKEVRPFNQLVPLISDPALINLRIVHVLVDPRAVRSREAGPLIA 220
 Db 162 MVEKAQRSHCFLVKEVRFLSQLQYLPLTDPSLNLYHVLVDPRAVRSREHTTELM 221

Qy 221 RDNGIVLTGIGKTYEADPHRLIRE-----VCRSHVRIAEA-ATLKPPFLRGY 269
 Db 222 VDSHVIG-----QHLETKEEDQPYAMKICKSQDIVKAIQL--PBAQORY 270

Qy 270 RLYRFEDLAREPLAIEIRALYAFGTGLTLPOLEAWIHNTHGSGKGKPIEAEHTSSRNARN 329
 Db 271 LFRLYEDLVRAPLAOTRYKEFGLDFLPHLOTVNVNTRKKGKQ--HAFTNARNALN 328

Qy 330 VSQWRHALFTKLRLRQEVCAGALOLQSYRPVYSAODQDLTLDLV 376
 Db 329 VSQAWRWSLPEYKSQLDQAGEMDLGYLQVRSQQEGNLSDLIL 375

RESULT 11

Q794G9 PRELIMINARY; PRT; 483 AA.

ID Q794G9 ID 088276 PRELIMINARY; PRT; 530 AA.

AC Q794G9; ID 088276; PRELIMINARY; PRT; 530 AA.

DT 05-JUL-2004 (TREMBrel. 27, Created)
 DT 05-JUL-2004 (TREMBrel. 27, Last sequence update)

DT 25-OCT-2004 (TREMBrel. 28, Last annotation update)
 DE N-acetylglucosamine-6-O-sulfotransferase (N-Acetylglucosamine 6-O-Sulfotransferase-1).
 DE Name=Gn6t-1;
 GN OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

| | | |
|-----------|---|---|
| RC | TISSUE=Whole embryo; | Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loqueline D.R., Abramson G.J., Peters G.J., McKernan K.J., Malek J.A., Gunaratne P.H., Bosak S.A., McEvlan P.J., Richards S., Worley K.C., Hale S., Garcia A.N., Gay L.J., Hulyk S.W., Villaon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madden A., Young A.C., Shevchenko Y., Bouffard G.G., J. Biol. Chem. 273:22577-22583 (1998). |
| RX | Mitsuoaka C., Kannagi R., Habuchi O., Muranatsu T.; "Molecular cloning and characterization of an N-acetylglucosamine-6-O-sulfotransferase"; J. Biol. Chem. 273:22577-22583 (1998). | Blakesley R.W., Touchman J.W., Grimwood J., Myers R.M., Butterfield Y.S., Rodriguez A.C., Krzywinski M.I., Skalska U., Smailus D.E., Schein A., Jones S.J., Scherich A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:1699-16903 (2002). [2] |
| RA | RA | RN |
| RA | Raha S.S., Loqueline D.R., Abramson G.J., Peters G.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.N., Gay L.J., Hulyk S.W., Villaon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madden A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Grimwood J., Myers R.M., Butterfield Y.S., Rodriguez A.C., Krzywinski M.I., Skalska U., Smailus D.E., Schein A., Jones S.J., Scherich A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:1699-16903 (2002). [2] | RN |
| RA | SEQUENCE FROM N.A. | RN |
| RA | STRAIN=C57BL/6; TISSUE=Brain; | RN |
| RA | Straubberg R.; Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases | RN |
| RA | EMBL; BC051963; AAH51963.1; -. | RN |
| RA | MSD; NCBI:1891167; Chet2. | RN |
| RA | CO; GO:0005615; C: extracellular space; TAS. | RN |
| RA | GO; GO:0006044; P:N-acetylglucosamine metabolism; TAS. | RN |
| RA | InterPro; IPR00863; Sulfotransferase. | RN |
| RA | Pfam; PF00685; Sulfotransferase_1. | RN |
| KW | Transferase. | RN |
| SQ | SEQUENCE 530 AA; 57814 MW; A113E1B735C363EC CRC4; | RN |
| Qy | Query Match 32.4%; Score 665; DB 2; Length 530; Best Local Similarity 40.4%; Pred. No. 8.9e-51; Matches 150; Conservative 64; Mismatches 121; Indels 36; Gaps 11; | RN |
| Db | 37 GGDEPRVHVLYVLSWRSGSSFLGOLFSQIDPVEYLMEPAWHWWTTLSGQAATHMAVDL 96 | DR |
| Qy | 160 GGDKRQLVYVFVTWRSGSSFLGELFQNPEVFTFLYEPWVWWQKLYPCDAVSLQGARDM 219 | DR |
| Db | 97 MRSFLCDMDVFDAYMPQ--SRNLs--A芬NATSAALSPPACSAFPRTISKQD--V 149 | DR |
| Db | 220 LSALYRCDLSVQFYSPAGSGGRNLTTLGIGATAKNTKVYCSSPLCPAYRKEVYGLVDRV 279 | DR |
| Qy | 150 CKTLCTROPSFLSARBCRSYSHVVLKEYRFLFNQVLYPLSDPALNLRIVHLVYDPAVL 209 | DR |
| Db | 280 CKK-CPPQRALARFEEBCRKYRTVIKGVRFDVAVAPLKDPLDKVHLVYDPAVA 338 | DR |
| Qy | 210 RS-----REAGPIIARD-----NGIVLT--NGKWTEDAPH-LRLIREV 246 | DR |
| Db | 339 SSRIRSRHGLIRESLQVRSRDPRAHMFPLFEAGHKIGAKKEGMGGPADAHYALGAMEVI 398 | DR |
| Qy | 247 CRSHVRIABAATLKKPPFLGRYVRLVPEDLAREPLAETRLAYAFTGTLTPOLEAMHN 306 | DR |
| Db | 399 CNSNAKTLOTA-LQPPDMIQGHVLYVVRFEDLVGPKTLRVRDFVGLVSEMEQFALN 457 | DR |
| Qy | 307 ITHGSG-IGKPIEAFTSSRNARNVSQAWRHALPFKTKLRYQEVCAAGALQQLGYRPVSA 365 | DR |
| Db | 458 MTSGSSSSKP---FVVSARNTAAANAWRALTFFQIQVEEFCYQPMAVGYERYNSP 514 | DR |
| Qy | 366 DQORDLTDLV 376 | DR |
| Db | 515 EEVRLDSKTL 525 | DR |
| RESULT 13 | | |
| Q80WV3 | PRELIMINARY; | PRT; 530 AA. |
| ID | SEQUENCE FROM N.A. | PRT; 483 AA. |
| AC | SEQUENCE=C57BL/6; TISSUE=Brain; | AC 09UED5, |
| DT | LINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; | DT 01-MAY-2000 (TREMBLrel. 13, Created) |
| RA | Straubberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Colins F.S., Wagner L., Shemmen C.M., Schulter G.D., Altschul S.F., Zeeberg B.B., Buettow K.H., Schaeffer M., Bhat N.K., Hopkins R.P., Jordan H., Moor T., Max S.I., Wang J., Haile F., Diatchenko L., Marusina K., Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. Name=Chst2; NCBI_TaxID=10090; DE Chst2_protein. | DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update) |
| OS | | DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update) |
| OX | | DE N-acetylglucosamine-6-O-sulfotransfase (GlcNAc6ST). |
| RN | | GN Name=SN6ST; |
| RC | | OS Homo sapiens (Human) |
| RA | | OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homino. |
| RA | | OC Mammalia; Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Stapleton M., Stapleton M., |

| | | | | | |
|--|--|----|--|--|--|
| OX | NCBI_TaxID=9606; | | | | |
| [1] | | | | | |
| RN | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE=Brain; | | | | |
| RX | MEDLINE=891845; PubMed=9722682; | | | | |
| RA | Uchimura K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q., Mitaoka C., Kannagi R., Habuchi O., Yokoyama I., Yamamura K., | | | | |
| RA | "Human N-acetylglicosamine-6-O-sulfotransferase involved in the biosynthesis of 6-sulfo sialyl Lewis X: molecular cloning, chromosomal mapping, and expression in various organs and tumor cells.", J. Biochem. 124:670-678(1998). | | | | |
| RN | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE=Placenta; | | | | |
| RX | MEDLINE=0498786; PubMed=11042394; DOI=10.1016/S0304-4165(00)00136-7; | | | | |
| RA | Sakaguchi H., Kitegawa H., Sugahara K.; "Functional expression and genomic structure of human N-acetylglucosamine-6-O-sulfotransferase that transfers sulfate to b-N-acetylglucosamine at the nonreducing end of an N-acetyllactosamine sequence.", Biochim. Biophys. Acta 1523:269-276(2000). | | | | |
| DR | EMBL; AB34679; BA34265; -; | | | | |
| DR | EMBL; AB21124; BA316886; -; | | | | |
| DR | EMBL; AB21125; BA316887; -; | | | | |
| DR | GO: GO:0016740; F:sulfotransferase activity; IEA. | | | | |
| DR | InterPro; IPR00086; Sulfotransferase activity; IEA. | | | | |
| DR | Pfam; PF00685; Sulfotransferase_1; 1. | | | | |
| KW | Transferase. | | | | |
| SQ | SEQUENCE 483 AA; 52787 MW; E27797D44931BA18 CRC64; | | | | |
| Query Match Score 659.5; DB 2; Length 483; Best Local Similarity 39.1%; Pred. No. 2.4e-50; Matches 151; Conservative 60; Mismatches 122; Indels 45; Gaps 12; | Query Match Score 659.5; DB 2; Length 530; Best Local Similarity 39.1%; Pred. No. 2.8e-50; Matches 151; Conservative 68; Mismatches 122; Indels 45; Gaps 12; | | | | |
| Qy | 31 GPSSPAG-----GEDRPHVLYLSSNSRGSSPFSGQLFSQHDPDVYLMEPAWHWTLL 81 | Qy | 31 GPSSPAG-----GEDRPHVLYLSSNSRGSSPFSGQLFSQHDPDVYLMEPAWHWTLL 81 | | |
| Db | 98 GVAAPGNGTRGTGEVGDKQVLVFTTSSGSSFFELINQNPAVEFFYLEPVWHWQRL 157 | Db | 145 GVAAPGNGTRGTGEVGDKQVLVFTTSSGSSFFELINQNPAVEFFYLEPVWHWQKL 204 | | |
| Qy | 82 SGQSATLHNAVRDLMRSIPLCDMDVFDAMQP---SRNLS--AFFNWATSRALCSPAC 136 | Qy | 82 SGQAATLHMAVRDLMRSIPLCDMDVFDAMQP---SRNLS--AFFNWATSRALCSPAC 136 | | |
| Db | 158 YPGDAVSLQGARDMLSAVRCDSVFOLYSPAGSGGRNNTTLGFGAAANTKVYCSPIC 217 | Db | 205 YPGDAVSLQGARDMLSAVRCDSVFOLYSPAGSGGRNNTTLGFGAAATNKVYCSPIC 264 | | |
| Qy | 137 SAFFPRTISKOD--VCKTLCTROPSSLAREACRSYSHVYLKEVRFENLQVLYPLISDPAL 194 | Qy | 137 SAFFPRTISKOD--VCKTLCTROPSSLAREACRSYSHVYLKEVRFENLQVLYPLISDPAL 194 | | |
| Db | 218 PAYRKEVVGLYDDRVCKK-CPPQRALARFEEBCKRTLVIGVRFDVAVLPLRDPAL 276 | Db | 265 PAYRKEVVGLYDDRVCKK-CPPQRALARFEEBCKRTLVIGVRFDVAVLPLRDPAL 323 | | |
| Qy | 195 NLRIVHLVYLPRAVRLS-----REAGPILARD-----NGIVLGNTNGKMW 234 | Qy | 195 NLRIVHLVYLPRAVRLS-----REAGPILARD-----NGIVLGNTNGKMW 234 | | |
| Db | 277 DLKTVHLVYLPRAVASSRIRSRRHGLJRESQVVRSDPRAHRMPFLEAAQHKLGAKKEGV 336 | Db | 324 DLKTVHLVYLPRAVASSRIRSRRHGLJRESQVVRSDPRAHRMPFLEAAQHKLGAKKEGV 383 | | |
| Qy | 235 --EADPH-LRLIREVCRSHVRIAETATLKPPPLGRYRILVRFEDLARPLAEIRALYAF 291 | Qy | 235 --EADPH-LRLIREVCRSHVRIAETATLKPPPLGRYRILVRFEDLARPLAEIRALYAF 291 | | |
| Db | 337 GGPDAYHALGAMEVICNSMANKLQTA-LQPPDWLQGHLYVRYSEBLVGDFVKTLRRYDF 395 | Db | 384 GGPDAYHALGAMEVICNSMANKLQTA-LQPPDWLQGHLYVRYSEBLVGDFVKTLRRYDF 442 | | |
| Qy | 292 TGLTLPOLBAWHNTTHGSG-IGKPIEAHTSSRNARYSQANRHLPPTKILRVQEV 350 | Qy | 292 TGLTLPOLBAWHNTTHGSG-IGKPIEAHTSSRNARYSQANRHLPPTKILRVQEV 350 | | |
| Db | 396 VGLLSPMEMPOALNTSGSSSSKR--FVVSARNATQANANWRTALTFQQIKOVEEFC 452 | Db | 443 VGLLSPMEMPOALNTSGSSSSKR--FVVSARNATQANANWRTALTFQQIKOVEEFC 499 | | |
| Qy | 351 AGALQLGYRPVYSAQDQRDLTDLV 376 | Qy | 351 AGALQLGYRPVYSAQDQRDLTDLV 376 | | |
| Db | 453 YQPMAVLGYRPVYSAQDQRDLTDLV 478 | Db | 500 YQPMAVLGYRPVYSAQDQRDLTDLV 525 | | |
| RESULT 15 | | | | | |
| Q9Y4C5 | PRELIMINARY; | | | | |
| ID | Q9Y4C5; OG9ZNS; Q9Y6F2; | | | | |
| AC | | | | | |
| DT | 01-NOV-1999 (T-EMBLrel. 12, Created) | | | | |
| DT | 01-MAR-2003 (T-EMBLrel. 23, Last sequence update) | | | | |
| DT | 05-JUL-2004 (T-EMBLrel. 27, Last annotation update) | | | | |
| DE | Long form of N-acetylglucosamine-6-O-sulfotransferase (GlcnAC6ST) | | | | |

Search completed: June 23, 2005, 08:48:58
Job time : 43.7985 sec

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - Protein search, using SW model

Run on: June 23, 2005, 08:33:38 ; Search time 8.8459 Seconds
 (without alignments)
 4242.027 Million cell updates/sec

Title: US-10-697-828-8

Perfect score: 2051

Sequence: 1 MWLPRFSSKTVTLLAQTT.....LTLDIVLPGRGPDHFSWASPD 390

Scoring table: BL05M62 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing first 45 summaries

Database : PIR_79;*
 1: Pir1;*
 2: Pir2;*
 3: Pir3;*
 4: Pir4;*

Searched:

Total number of hits satisfying chosen parameters:

283416

RESULT 1

JE0261

N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human

C;Species: Homo sapiens (man)

C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 01-Mar-2002

C;Accession: JE0261

R.Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Mitaoka, T.

J. Biochem. 124, 670-678, 1998

A;Reference number: JE0261; PMID: 98391845; PMID: 9722682

A;Accession: JE0261

A;Molecule type: mRNA

A;Residues: 1-484 <UCH>

A;Cross-references: DDBJ:AB014679

C;Comment: This protein catalyzes the transfer of sulfate from 3'-phosphoadenosine to the 6-OH group of N-acetylglucosamine.

C;Superfamily: chondroitin 6-sulfotransferase

C;Keywords: sulfotransferase

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|------------------------|
| 1 | 659.5 | 32.2 | 484 | 2 JE0261 | N-acetylglucosamin |
| 2 | 613.5 | 29.9 | 484 | 2 JC7350 | N-acetylglucosamin |
| 3 | 610.5 | 29.7 | 486 | 2 JC7351 | N-acetylglucosamin |
| 4 | 540.5 | 26.4 | 458 | 2 A57397 | chondroitin 6-sulf |
| 5 | 107.5 | 5.2 | 307 | 2 B95934 | probable enzyme, C |
| 6 | 102.5 | 5.1 | 359 | 2 T16350 | hypothetical prote |
| 7 | 102.5 | 5.0 | 388 | 2 G70729 | pyruvate dehydroge |
| 8 | 96 | 4.7 | 388 | 1 DEHUPT | phenylalanine amino |
| 9 | 96 | 4.7 | 707 | 2 S29029 | probable valine-py |
| 10 | 95 | 4.6 | 474 | 2 T38909 | endopeptidase Clp |
| 11 | 95 | 4.6 | 885 | 1 S78246 | peptidomeric glycoprot |
| 12 | 94.5 | 4.6 | 1581 | 1 GJWJBV | probable oxalyl-Co |
| 13 | 94 | 4.6 | 582 | 2 A70841 | DNA polymerase III |
| 14 | 93.5 | 4.6 | 615 | 2 C75278 | probable ATP-bind |
| 15 | 93 | 4.5 | 523 | 2 B83629 | tyrosylprotein sul |
| 16 | 90.5 | 4.4 | 380 | 2 T42755 | hypothetical prote |
| 17 | 88 | 4.3 | 737 | 2 AD2037 | hypothetical prote |
| 18 | 86.5 | 4.2 | 445 | 2 T05639 | probable long-chai |
| 19 | 86.5 | 4.2 | 577 | 2 B75585 | hypothetical prote |
| 20 | 86.5 | 4.2 | 712 | 2 T33231 | hypothetical prote |
| 21 | 86 | 4.2 | 417 | 2 F12238 | hypothetical prote |
| 22 | 85.5 | 4.2 | 221 | 2 T50665 | pyruvate dehydroge |
| 23 | 85.5 | 4.2 | 389 | 1 DEPGPA | 3-deoxy-D-manno-oc |
| 24 | 85.5 | 4.2 | 425 | 2 B83023 | rotein kinase cak |
| 25 | 85.5 | 4.2 | 571 | 2 S69210 | mutS family DNA mi |
| 26 | 85.5 | 4.2 | 1254 | 2 T41262 | cholesterol hydr |
| 27 | 85 | 4.1 | 364 | 2 A49781 | hypothetical prote |
| 28 | 85 | 4.1 | 696 | 2 AI1029 | membrane protein t |
| 29 | 85 | 4.1 | 776 | 2 T31210 | |

ALIGNMENTS

| RESULT 2 | | RESULT 3 | |
|--|---------|--|---------|
| C7350 | <UCH> | C7351 | <UCH> |
| -acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] | - mouse | -acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] | - human |
| :Species: Mus musculus (house mouse) | | :Species: Homo sapiens (man) | |
| :Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004 | | :Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004 | |
| :Accession: JC7350 | | :Accession: JC7351 | |
| :Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in little. Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a | | :Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in little. Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a | |
| :Keywords: glycosylation; sulfotransferase; sulfate transport; transmembrane protein | | :Keywords: glycosylation; sulfotransferase; sulfate transport; transmembrane protein | |
| :Reference number: JC7350; MUID:20374462; PMID:1091333 | | :Reference number: JC7351; MUID:20374462; PMID:1091333 | |
| :Accession: JC7350 | | :Accession: JC7351 | |
| :Molecule type: mRNA | | :Molecule type: mRNA | |
| :Residues: 1-484 <UCH> | | :Residues: 1-486 <UCH> | |
| :Cross-references: UNIPROT:Q99NBO; UNIPROT:Q9EP78; DDBJ:AB040710 | | :Cross-references: UNIPROT:Q9EP78; DDBJ:AB040710 | |
| :Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in little. Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a | | :Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in little. Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a | |
| :Keywords: glycosylation; sulfotransferase; sulfate transport; transmembrane protein | | :Keywords: glycosylation; sulfotransferase; sulfate transport; transmembrane protein | |
| Query Match Score 613.5; DB 2; Length 484; | | Query Match Score 610; DB 2; Length 486; | |
| Best Local Similarity 38.8%; Pred. No. 2.1e-47; | | Best Local Similarity 38.8%; Pred. No. 4.4e-11; | |
| Matches 146; Conservative 54; Mismatches 127; Indels 49; Gaps 11; | | Matches 147; Conservative 52; Mismatches 126; Indels 54; Gaps 11; | |
| 30 PGPPSPAG---GEDRVHVLYLSSWRSGSSPFLQLFSQHPDFYLMEPANHWTITLSQGSA 86 | | 30 PGPPSPAG---GEDRVHVLYLSSWRSGSSPFLQLFSQHPDFYLMEPANHWTITLSQGSA 86 | |
| 85 PGNLDAVGEAVTQEQHIIYVHAIWGSSFLIGELLNQHPDFYLMEPWMHLWQAIYPGDA 144 | | 85 PGNLDAVGEAVTQEQHIIYVHAIWGSSFLIGELLNQHPDFYLMEPWMHLWQAIYPGDA 144 | |
| 87 ATLHNAYRDLMRSTIFCLDMDFVFDAY-----MFQPSRNLSA-FENWATSRALCSPPA 135 | | 87 ATLHNAYRDLMRSTIFCLDMDFVFDAY-----MFQPSRNLSA-FENWATSRALCSPPA 135 | |
| 145 ESQLQALDRDMLRSFLRCDFSVLRLAQPGDGERAFPSANLTAMLFWRNTKVCISPPPL 204 | | 145 ESQLQALDRDMLRSFLRCDFSVLRLAQPGDGERAFPSANLTAMLFWRNTKVCISPPPL 204 | |
| 136 CSAFRGTT---ISQDVCKTLCTQFDSL AREA CRSYSHVVKKEVRFVNLOVLYPLL 189 | | 136 CSAFRGTT---ISQDVCKTLCTQFDSL AREA CRSYSHVVKKEVRFVNLOVLYPLL 189 | |
| 205 CPAARADAVGVLVEDKACESTC--PPVSLRALEECRKV PVVVTKDVRLLDLGTVLPULL 262 | | 205 CPAARADAVGVLVEDKACESTC--PPVSLRALEECRKV PVVVTKDVRLLDLGTVLPULL 262 | |
| 190 SDPAANLRIVHLVRDPRAVRSREAGAPILARDNGIVLGTINGK-----WVEADP 238 | | 190 SDPAANLRIVHLVRDPRAVRSREAGAPILARDNGIVLGTINGK-----WVEADP 238 | |
| 263 RDPGMLNKVQFLFRDRAVNSRLKSQGLLRESQVLRPQRGDFHRVILLAHGVDARP 322 | | 263 RDPGMLNKVQFLFRDRAVNSRLKSQGLLRESQVLRPQRGDFHRVILLAHGVDARP 322 | |
| 239 -HFLRLIREVCRSHYRIAEEA-----TUKKPPFLRGYRFLVRFEDLAIREPLA 283 | | 239 -HFLRLIREVCRSHYRIAEEA-----TUKKPPFLRGYRFLVRFEDLAIREPLA 283 | |
| 323 GGQAQALPSAPRADFLTSALLEVICAEAWLDRDFFTRGAPEAWLRRRLRKYEDLVWPQQA 382 | | 323 GGQAQALPSAPRADFLTSALLEVICAEAWLDRDFFTRGAPEAWLRRRLRKYEDLVWPQQA 382 | |
| 284 EIRALYAFTGILTTLPOLEAHWNITHGSGICPKPIEAFTSSRNAAVNSQAMRHAPFTKI 343 | | 284 EIRALYAFTGILTTLPOLEAHWNITHGSGICPKPIEAFTSSRNAAVNSQAMRHAPFTKI 343 | |
| 383 QLRRLRFSGURTLAALDAFAFNTRGSAYGAD-REFHLSARDAREAVHWRERLSQEQV 441 | | 383 QLRRLRFSGURTLAALDAFAFNTRGSAYGAD-REFHLSARDAREAVHWRERLSQEQV 441 | |
| 344 LRVQAVCAGALQLOGY 359 | | 344 LRVQAVCAGALQLOGY 359 | |
| 442 RQVETACAPMRLLAY 457 | | 442 RQVETACAPMRLLAY 457 | |

Db 393 IYSTQKNSSEQQPEKWRPSIPEKLAQVYDQACEPAMRLFGYKLASSAQBLTNRSLSL-LBEE 451
 Qy 380 GP 381
 Db 452 GP 453

RESULT 5

B95934 probable enzyme, C-terminal domain similar to sulfotransferase protein [imported] - Sinc
 C;Species: Sinorhizobium meliloti
 C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C;Accession: B95934
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A;Title: The composite sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endoc
 A;Reference number: A95842; MUID:21396508; PMID:11481431
 A;Accession: B95934
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-307 <KUR>
 A;Cross-references: UNIPROT:Q92VG4; GB:AL591985; PIDN:CAC49141.1; PID:g15140626; GSPDB:G
 A;Experimental source: strain 102+, megaplasmid pSymB
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 L.; Chaix, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 Science 293, 668-672, 2001
 A;Authors: Kahn, D.; Kalman, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lealaure,
 hebault, P.; Vandembol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A;Reference number: A96039; MUID:21368234; PMID:11474104
 A;Contents: annotation
 C;Genetics:
 A;Gene: SMB21237
 A;Genome: plasmid

Query Match 5.2%; Score 107.5; DB 2; Length 307;
 Best Local Similarity 20.6%; Pred. No. 0.064;
 Matches 78; Conservative 46; Mismatches 131; Indels 123; Gaps 17;

Db 32 PSSPAGGERVHVLVLLSW-RSGSSFLQQLFSQHDPVF--YLMBPWVHWTLSQGSA 87
 Db 2 PSQP-----VRIAYIGYGRGSTLIDALGQHAHVAGEITSLTRHW---RHNEYC 52
 Qy 88 TLHMAVRDLRSIPLCD-----MDVFDYMPQSRNLSTAFFWATSRLA 130
 Db 53 ACGNAIRD-----CSFWSSVRRENSDGGDQGLMEYCALQKEFGLSMTRLLSGMGL 105
 Qy 131 CSPPACSAFPRTISKQDVCKTLCPROPFLAREACRSYSHVLFKEVRFNLQVLYPLIS 190
 Db 106 -----GQFSLYLHPTKRLFS-AMQSC-SGRQVIVDSSKLFGRAMAQAII 148
 Qy 191 DPALNLRIVHLVDRPAVLRSREAAGPILLARDNGIVLGTMNKWEDAPHHLIREVCRSH 250
 Db 149 -PGIDMRVHLVRDGRGV-----AVSLL-----KGYERDAKGLQKBI---- 185
 Qy 251 VRIAEEATLKEPPFLGRYR-----LYRVEDLAREPLAEIRALY 289
 Db 186 -----KPKSVFTALRWSMVNLAVEYLSRKLGSEKVNRYVVEDADSDPVAMMQIG 236
 Qy 290 AFTGILUTPQLEBAWIIHNTSSGIGKPIAFHTSSRN-----ARNVSOAWHLPFT 341
 Db 237 TPLELDLS-QVGTSLEN-----GEAMGPGHQVAGNRUFLRMNASTALKDETWRTRMPAR 288
 Qy 342 KILRVQEBCAGLQLGY 359
 Db 289 QVSFQLGWMRLRGY 306

C;Species: Caenorhabditis elegans
 C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 R;Raich, A.
 A;Submitted to the EMBL Data Library, March 1996
 A;Description: The sequence of *C. elegans* cosmid F42G9.
 A;Accession: T16350
 A;Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residues: 1-359 <TAI>
 A;Cross-references: UNIPROT:Q20351; EMBL:U00051; NID:91216305; PID:91216308; PIDN:AAA9135
 A;Experimental source: strain Bristol N2
 C;Genetics:
 A;Gene: CESP:F42G9.8
 A;Introns: 28/2; 50/2; 113/2; 155/3; 220/3; 290/2
 C;Superfamily: Caenorhabditis elegans hypothetical protein F42G9.8
 Query Match 5.1%; Score 105; DB 2; Length 359;
 Best Local Similarity 19.8%; Pred. No. 0.13;
 Matches 80; Conservative 53; Mismatches 131; Indels 140; Gaps 21;

Db 11 VTVLLAQTCCLLFI-----ISRPGPSPAGGEDBDRVH-----VLY 46
 Db 31 YYIFIFCFICLICLIFSTIKCKLQEKLEQLSLSKESLNEODARHSRRLSNSDEQQLIFV 90
 Qy 47 LSSWRSGSSFLQGLFSQHDFV-----FVLMEPW-HVWTLSQGSAATHMAVR 94
 Db 91 GGVRSGTTLMRALDAHDVRCGGETMLLPSLTWQAGWRNDWV---NNSGIT---- 141
 Qy 95 DLMSRIFLCDDMDVFDAYMPQSRNLUFFWATSRLCSPACSAFPRTISK-QDVCCKTL 153
 Db 142 -----QEVD-----DAVASAPTEIVAKHSBAPRL 167
 Qy 154 CTQPFSLAREACRSYSHVLFKEYRFPNLQVLYPLISDPALNLRIVHLYDPRAVLRSR- 212
 Db 168 CNKD-----YTALWLPTTR-----RLYP-----NAKFPLMIRDARA7VHSMI 205
 Qy 213 EAAGPILARDNGIVLGTNGKWEADPHFLIREVCRSHVRIAETATLKPPFLGRYRLV 272
 Db 206 ERKVPAVGTNTSDEISMPVQW---NOELRQMTFCQNN---APGQCIK-----V 247
 Qy 273 RFEDLAREPLAEIRALYAFGTLTLPQLEBAWIIHNTSSGIGKPI----EAFTHS-S-RNA 327
 Db 248 YIERLJQKPAEIRIRTNFLDLPSPQM-----LRHQDLIGDEVLDNDOEFSAAQVKNS 301
 Qy 328 RNVS--QAWHALPPTKILRQEVCAGALQLGY----RPVYS 364
 Db 303 INTKALTNSFDCCFSEETURKLDVV-AFFLGILGDTTSISKPDYs 344

RESULT 7
 G70729 hypothetical protein Rv2267c - Mycobacterium tuberculosis (strain H37RV)
 C;Species: Mycobacterium tuberculosis
 C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C;Accession: G70729
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.B.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome &
 A;Molecule type: DNA
 A;Residues: 1-388 <COL>
 A;Cross-references: UNIPROT:050695; GB:277163; NID:93261610; PIDN:CA00968.1
 C;Genetics:
 A;Gene: Rv2267c

RESULT 6
 T16350 hypothetical protein F42G9.8 - Caenorhabditis elegans

| | | | | | |
|--|--|-----------------|----------------|-------------|---|
| Query Match | 5.0% | Score 102.5; | DB 2; | Length 388; | |
| Best Local Similarity | 22.7% | Pred. No. 0.25; | Mismatches 37; | Indels 89; | Gaps 15; |
| Matches 66; | Conservative | No. 0.25; | Mismatches 99; | Indels 89; | Gaps 15; |
| Qy | 44 VLVLSWRSSSFIGOLF--SOH-----PDVFYLMPEAWHYWTTSQGSAAITLH 90 | | | | RESULT 9 |
| Db | 83 IPIVGHWRTGLTLLHELYVDDRHTGPGYBCLAPHHLTLE----WFAPYVEFLVSKH 137 | | | | phenylalanine ammonia-lyase (EC 4.3.1.5) - sweet potato |
| C | 91 MAYDLIMSLFICLMDVFDAYMPOQRNISAFFWATSAICSPACSAFRPTISKDVC 150 | | | | C.Species: Ipomoea batatas (sweet potato) |
| C | 138 RANDNMDSL-----HHQDEDE---FWWC-HQGLPSVLTIAFP----- 172 | | | | C.Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004 |
| C | 151 KTLCTROP----FSLAREA CRSYSHVVLKEFVNLOVY----PLSDPALNRI- 198 | | | | C.Accession: S29029 |
| C | 173 ---NRBPQEYEDLQVAPRLE--IWKRTLTFEVQVYPRRKTVILKOPTHSRIK 226 | | | | A.Title: Structure and characterization of a cDNA clone for phenylalanine ammonia-lyase 1 |
| C | 199 -----VHLVRDPRAVRSREAAQPIIARDNGIVL----GTNCKWVADPHHLRLI 243 | | | | A.Reference number: S29029 |
| C | 227 VLLVEFPQAKFHVLRDPVVYVSTIHLRALKRIGQPTDGLDJKVYST--YVDY 284 | | | | A.Molecule type: mRNA |
| C | 244 REVERSHYRIAENATLKKPFPFLGRYLVRFEDLAREPLAEIRALYARTGL 294 | | | | A.Residues: 1-707 <TRAN> |
| C | 285 R-----KLDEGRELVDEP----RFYELRYEDLIGDPEGGQLRRYQHGLT 324 | | | | A.Cross-references: UNIPROT:P14166; EMBL:M29322; PIDN:AAA13389_1; PID:gi16821; |
| Qy | 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 | | | | C.Superfamily: histidine ammonia-lyase |
| Db | 131 FGNATECSHTLPHSATAAAMLYRINTLLOGYSGIVLTTKLNNTCPLP-R-G 188 | | | | C.Keywords: ammonia-lyase; carbon-nitrogen lyase |
| Db | 201 LYRDPRAVLRSREAAQPIIARDNGIVLGTNGKRVEADHHLRLIRECRSHYRIEA 257 | | | | F.192-194/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted |
| Db | 189 TTASGDIVPLSYAGLITGRNSKAIGPNTLNAEALRL---SRSGRRIFRVASPR 244 | | | | F.193/Modified site: dehydroalanine (Ser) #status predicted |
| Qy | 258 ---TKPPPFYLRGRYLRVRFEDLAREPLAFT-RALYA-----PTGLTL-TPOE--- 301 | | | | Query Match |
| Db | 245 KGJPSLNUAPPLYGMANVLFANVLPVLSATAEVNGKPEPTD-HLTHKLXHHP 303 | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| Db | 302 -----AMHNITHSGSGIGKPLEAFH-----TSSRNARNVSQAW 334 | | | | Matches 57; Conservative |
| Db | 304 QTEAAATIMEHLDGSSSYVKAQKLUHEMDPLQKPKQRYALRTSPQW 350 | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| RESULT 10 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| Db | 378905 T38905 probable valine-pyruvate transaminase (EC 2.6.1.66) SPAC56B4_03 [similarity] - fission yeast | | | | Matches 57; Conservative |
| C | Alternate names: pyruvate dehydrogenase complex, E1 component alpha chain | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| C | Species: Homo sapiens (man) | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| C | Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004 | | | | Matches 57; Conservative |
| C | Accession: A7104 | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| R.Dahl, H.H.M.; Brown, R.M.; Hutchison, W.M.; Maragos, C.; Brown, G.K. | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| Genomics 8, 225-232, 1990 | | | | | Matches 57; Conservative |
| A.Title: A testis-specific form of the human pyruvate dehydrogenase Elalpha subunit is c | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A.Reference number: A37104; MUID:91065637; PMID:2249846 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A.Accession: A7104 | | | | | Matches 57; Conservative |
| A.Molecule type: mRNA | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A.Residues: 1-388 <DAHs> | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A.Cross-references: UNIPROT:P29803; GB:M86808; GB:J04769; PIDN:AAA60232_1; C;Genetics: | | | | | Matches 57; Conservative |
| A;Gene: GDB:PDHA2 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Cross-references: GDB:120711; OMIM:179061 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 4q22-4q23 | | | | | Matches 57; Conservative |
| C;Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain: thiamin pyrophosphate-binding flavoprotein; heterotrimer; mitochondrial; oxidoreductase; phosphoprotein; F.27/Domain: transit peptide (mitochondrion) #status predicted <TNP> | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| F.28-388/Product: pyruvate dehydrogenase (lipoamide) alpha chain #status predicted <MAT> F.18-232/Domain: thiamin pyrophosphate-binding domain homology <TPB> | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| F.291/Binding site: phosphate (Ser) (covalent) #status predicted F.298/Binding site: phosphate (Ser) (covalent) #status predicted | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Residues: 1-474 <CON> | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Cross-references: UNIPROT:O14192; EMBL:Z99261; PID:gi4038621; PIDN:CA816394_1; PID:g2414; C;Generics: | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: Escherichia coli valine-pyruvate transaminase C;Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1-474 <CON> | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Status: submitted to the EMBL Data Library, September 1997 | | | | | Matches 57; Conservative |
| A;Reference number: 221813 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Accession: T38905 | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Molecule type: DNA | | | | | Matches 57; Conservative |
| A;Experiment source: strain 972h-; cosmid c56E4 | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |
| A;Status: translated from GB/EMBL/DBDJ | | | | | Best Local Similarity 25.1%; Pred. No. 2.1%; Mismatches 26; Indels 86; Gaps 10; |
| A;Map position: 1 | | | | | Matches 57; Conservative |
| C;Superfamily: aminotransferase; phosphoprotein; pyridoxal phosphate F.297/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted | | | | | Qy 159 FSLAREACRCSYSHVVLKEFVNLOVYPLIS--DPALNIRIVH 200 |

| | | | | | |
|----|--|-----|----|--|-----|
| Db | 128 NIIMTTGNTSCLDIALRMLTNRGDSILVEKYSSFPSALQSMRPLGLSCIPIDMDQF-GFLP | 186 | Db | 211 KNPVPVLTGEPEGV-----GKTAVERGLAQI-----IAB-----KAPDFLDGNL---- | 248 |
| Qy | 114 QSRNLSAFFNW-ATSRALCSPACSAFPRTISIQDVCKTLCTRQPFSIAREACRSYHV | 172 | Qy | 274 FEDAREPAAETALAYATGTLTPOQEAWINITHGGIGKIEAFT-----SSRVAR | 328 |
| Db | 187 ESMDD-DILNTWDATSYGSPPRPHVLYTIPQNPNPGSTLSVERRQIYTIAQK---HDII | 241 | Db | 249 --LMALDQGSLIAGTKYRG-EFEEIRKRIIVEVQNSAIIVIDEHTLVGAGAAEGAV | 304 |
| Qy | 173 VLKEVRFENIQLQV-LY---PLSDPDALNLRVHLYRDPRAWLRSREAAQPLARDN---GI | 225 | Qy | 329 NVSQAWRHALPFTKILRYQEVCAQALQLGYR-----PVYSDAQORDLTLD | 374 |
| Db | 242 ILEDDPYYKPKDADKAFTNE-QFYLRLISFLSDVDRVRMDSLSKV | 239 | Qy | 305 DAANTLKDPLARSKFR-----CIGATITDEYRKYIERDPALERRFQPVHVKPEPTVGVTIE | 359 |
| Qy | 226 VLGNGKWTBADPHLRLIRECRSHVRIAATLKPPPLRG-RYRLYRF--EDLAREPL | 282 | Qy | 375 LVL 377 | |
| Db | 300 APGRSGWGFPAQPLF----IEGLRLRAETATQSASGISQGILYAMFHMGQDGYLEWL | 353 | Db | 360 ILL 362 | |
| Qy | 283 AEIRALYAFF | 292 | | | |
| Db | 354 KHIR-YSTF | 361 | | | |
| | RESULT 12 | | | | |
| | VGNQBV | | | | |
| | peplomer glycoprotein precursor - Berne virus (strain P138/72) | | | | |
| | C;Species: Berne virus | | | | |
| | C;Accession: A36759 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004 | | | | |
| | C;Accession: A36759 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004 | | | | |
| | N;Alternate name: ATP-dependent Clp protease regulatory chain; caseinolytic Clp protease; adenosine triphosphatase (EC 3.6.1.3) | | | | |
| | C;Species: chloroplast Odontella sinensis | | | | |
| | C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004 | | | | |
| | C;Accession: S78246 | | | | |
| | R;Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U. | | | | |
| | Plant Mol. Biol. Rep. 13, 336-342, 1995 | | | | |
| | A;Title: The Chloroplast Genome of a chlorophyll a+c-containing Alga, Odontella sinensis | | | | |
| | A;Reference number: S78246 | | | | |
| | A;Status: nucleic acid sequence not shown; translation not shown | | | | |
| | A;Molecule type: DNA | | | | |
| | A;Function: allows clipp to hydrolyze polypeptides and proteins, probably by a chaperone activity; ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller polypeptides; endopeptidase Clp ATP-binding chain | | | | |
| | C;Superfamily: nucleotide-binding motif A (P-loop) | | | | |
| | C;Keywords: ATP; chloroplast; hydrolase; molecular chaperone; nucleotide binding; P-loop | | | | |
| | F;248-225;Region: nucleotide-binding motif A (P-loop) | | | | |
| | F;50-567;Region: nucleotide-binding motif B | | | | |
| | F;628-633;Region: nucleotide-binding motif B | | | | |
| | F;524/Binding site: ATP (Lys) #status predicted | | | | |
| | F;566/Binding site: ATP (Lys) #status predicted | | | | |
| | A;Gene: clpc | | | | |
| | A;Genome: chloroplast | | | | |
| | C;Function: | | | | |
| | A;Description: allows clipp to hydrolyze polypeptides and proteins, probably by a chaperone activity; ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller polypeptides; endopeptidase Clp ATP-binding chain | | | | |
| | C;Superfamily: nucleotide-binding motif A (P-loop) | | | | |
| | C;Keywords: ATP; chloroplast; hydrolase; molecular chaperone; nucleotide binding; P-loop | | | | |
| | F;248-225;Region: nucleotide-binding motif A (P-loop) | | | | |
| | F;50-567;Region: nucleotide-binding motif B | | | | |
| | F;628-633;Region: nucleotide-binding motif B | | | | |
| | F;524/Binding site: ATP (Lys) #status predicted | | | | |
| | F;566/Binding site: ATP (Lys) #status predicted | | | | |
| | Query Match 4.6%; Score 95; DB 1; Length 885; | | | | |
| | Best Local Similarity 21.0%; Pred. No. 3.5%; Mismatches 59; Indels 114; Gaps 18; | | | | |
| | Matches 89; Conservative 59; N mismatches 161; Delins 114; Gaps 18; | | | | |
| Qy | 5 RFSSTKTVLILLAQ-----TCUCLLFI-SRPGPSSPAGGEDRHYVLVLSMR 51 | | Qy | 181 NL-----QVLYPL-----LSDPALNL-----RIVHLVDRPRAVL-----209 | |
| Db | 4 KPTEGAIVKVMLSQEARRMGHNFVGTQEJLJGIGQRHIGARALKKQV---TLKKR | 60 | Db | 361 TLFSTLFSMDYQINPQLQVLTQNETFLQLPATNYQGOTLISQGRMLNFKDIAVFLDFD | 420 |
| Qy | 52 -----SGSSSELGOLFQHPDVYLMPEAWHMTLTSQSGSAATHMARDLMSR-1F 101 | | Qy | 474 HP---AGVINIARHHSMSVAARNSTSFCYCVTHSLTTFGKLIDISTSWFFHTLALPSGPVSQ | 531 |
| Db | 61 REIETYIGRTGFGEVASEIPTPRAKRVLLENAVHEGKDGLGNNFGVTEHILALISESDGYA | 120 | Db | 210 ---RSREAGP---ILLARDNGTIVLGLTGNGKWWBAPDHFLIRECRSHVRIAATLKP | 261 |
| Qy | 102 LCDMDVFDAMPOQSNLSAFPENWATSAACSPACSAFPRTISKODVCKTLCTRQPFSI | 161 | Db | 421 TKFVRTNDPSSDIFVUVVARQAQLIYGNFRIQBINGFQV--KCSSNI---ISTLEP | 473 |
| Db | 121 MRTLDKLGVNIPKLNLLILAYI-----EENOEBELRPLTQAEEKFL | 161 | Db | 262 PPFLRGRLVRFEDLAREPLAETRLAFT----- | 292 |
| Qy | 162 AREACRS-----YSHVVLKEVRFENLQVLYPLISDPAIHL-RIVHLVDRPRAVLRSRE | 213 | Qy | 474 HP---AGVINIARHHSMSVAARNSTSFCYCVTHSLTTFGKLIDISTSWFFHTLALPSGPVSQ | 531 |
| Db | 162 EREKGGSSTPLDESENISKEAVDGKL-----DPVIGRDKEIHEV--IKVLLARR | 210 | Db | 293 -----GLITLPQLEAHINNT-----HGSGIGKPIEAFTHTSSRNARYS | 331 |
| Qy | 214 AAGPILARDINGIVLGTNGKWEADPHLRLIRECRSHVRAEAATLKPPFLGRYLYR | 273 | Db | 532 VSMPLIPLSTAAGVYMMHPMEIHWIPLLTLAQSQTOPSFFIGINKTI-TLTTOLOAYAQVY | 590 |
| | | | Qy | 332 QAWRHALPFTKILRVQEVCAQALQLGYRPFVYSAADQRDL | 371 |

| | | | |
|-----------|--|---|---|
| Db | 591 TAWFLSVIYVRLPARRLTG-YQLVVPFTIQALSIKQADL 629 | RESULT 13 | Query Match 4.6%; Score 93.5; DB 2; Length 615; Best Local Similarity 23.4%; Pred. No. 2.9; Matches 92; Conservative 41; Mismatches 141; Indels 119; Gaps 21; |
| A7041 | decarboxylase - Mycobacterium tuberculosis (strain H37RV) | C:Species: Mycobacterium tuberculosis C:Accession: A7041 R.; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Ratnareddam, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. | Qy 4 PRFSSKTVTUVLIAORTCLLIFTISRGPSSPAGGEDRVHVLVLSWRSGSSFLGQFLFSQ 63 Db 44 PRGVKTTARLAMTA-----NCGPAPKPCBESCLAV----RAGS----- 83 |
| A7041 | #sequence_revision 17-Jul-1998 #text_change 16-Aug-2004 | R.; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Ratnareddam, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. | Qy 64 HDPVYLMPEPAHWTWTLSQSAATHMA--VRDLMSIFLCMD----VFDAYMPOS 115 Db 84 HDPMVMEI-----DAASNNNSVDDYDRERKVLGAAMRGCKKIYLDEAHMMS 129 |
| A7041 | Nature 393, 537-544, 1998 | Nature 393, 537-544, 1998 | Qy 116 RNLSAFPNWATSRALCCPPACSAFPRTISKDVKLCTROPSLAREACSY--SHVV 173 Db 130 R---AATN-ALKTLEPPHEVIFILATEPEKIPPLSLR-----COHYRFRRLT 176 |
| A7041 | Authors: Spores, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. | A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Reference number: A70500; MUID:98295987; PMID:9632320 | Qy 174 LKEV--REFENQVLYPLVPLSDP-ALNL--RIVH-LVRDPRAVLRSREAAAGPTLARDNGIVL 227 Db 177 SEIAGKLGLVTLLEGASADPDALNIGRLADGMARDGESLLERMLAGTAVTRP---- 231 |
| A7041 | A;Status: preliminary; nucleic acid sequence not shown; translation not shown | A;Status: preliminary; nucleic acid sequence not shown; translation not shown | Qy 178 LY--AFTGLLTTPOLEWIHNITHGS-GIGRJEAFTTSRNARNVQAWHALPFTKIL 344 Db 232 -----AVEALGP LPGEVNRGVASALIYGD-AEBAISGAAQ 266 |
| A7041 | A;Molecule type: DNA | A;Molecule type: DNA | Qy 228 GTNGKWTeadPHURLIREVCRSHVRLAATLKPPPLURGVRYLVRFEDLAREPLAIBRA 287 Db 267 LYRDGFEARTVVGLVAFGAHLAEGLGE----- 310 |
| A7041 | A;Cross-references: UNIPROT:Q53639; GB:AL021926; PMID:93261520; PIDN:CAA1731 | A;Experimental source: strain H37RV | Qy 288 LY--AFTGLLTTPOLEWIHNITHGS-GIGRJEAFTTSRNARNVQAWHALPFTKIL 344 Db 311 KLQ---AALDEQEAPARSADQQ--SLELA 336 |
| RESULT 15 | | | |
| B83629 | probable ATP-binding component of ABC transporter PA0136 [Imported] - Pseudomonas aeruginosa | Pseudomonas aeruginosa | B83629 C:Accession: B83629 Ri-Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Brzustowski, T.; Yuan, Y.; Brody, L.R.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V. |
| B83629 | Sequence revision 15-Sep-2000 #text_change 09-Jul-2004 | | Nature 406, 959-964, 2000 |
| B83629 | A;Title: Complete Genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen | A;Reference number: B83629 | A;Cross-references: UNIPROT:Q91626; PMID:20437337; PMID:10984043 |
| B83629 | A;Status: Preliminary | A;Molecule type: DNA | A;Cross-references: 1-523 <STO> |
| B83629 | A;Genes: PA0136 | A;Residues: 1-523 | A;Experimental source: strain PA01 |
| RESULT 16 | | | |
| Db | 57 RHEASAGVAAAGCFLTARP-GVCLTTSGPGLNGLPALARATTNCFPMIQSGSSS-RP 114 | C:Species: Pseudomonas aeruginosa | Query Match 4.5%; Score 93; DB 2; Length 523; Best Local Similarity 26.4%; Pred. No. 2.6; Matches 37; Conservative 17; Mismatches 60; Indels 26; Gaps 5; |
| Db | 57 RHEASAGVAAAGCFLTARP-GVCLTTSGPGLNGLPALARATTNCFPMIQSGSSS-RP 114 | C:Accession: C75278 R.; White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M. | Qy 13 VILLAQTCLLIFTISRGPSSPAGGEDRVHVLVLSWRSGSSFLGOLFQHPDVFTLME 72 Db 384 VIALAE-BIIRPAVKAGAGPAPS-----LSGNLQKFILGEILQAPRLLVAH 434 |
| Db | 57 RHEASAGVAAAGCFLTARP-GVCLTTSGPGLNGLPALARATTNCFPMIQSGSSS-RP 114 | Science 286, 1571-1577, 1999 | Qy 73 PAWHWVTTLSQSAATHH--MAVRDLMRSFLCDMVFYDAMPOQSNLSAFENWATRSA 129 Db 435 PTWGV----DVAEALIHRALDRGATIAVVESELDLFLLSDR-----IAA 480 |
| Db | 57 RHEASAGVAAAGCFLTARP-GVCLTTSGPGLNGLPALARATTNCFPMIQSGSSS-RP 114 | A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. | Qy 130 LCSPPACSAFPRTISKQDV 149 Db 481 LCSGRUICPAVATASASSQKV 500 |
| RESULT 17 | | | |
| C75278 | DNA polymerase III, tau/gamma subunit - Deinococcus radiodurans (strain R1) | DNA polymerase III, tau/gamma subunit - Deinococcus radiodurans | C:Species: Deinococcus radiodurans C:Accession: C75278 R.; White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M. |
| C75278 | #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004 | Science 286, 1571-1577, 1999 | Qy 13 VILLAQTCLLIFTISRGPSSPAGGEDRVHVLVLSWRSGSSFLGOLFQHPDVFTLME 72 Db 384 VIALAE-BIIRPAVKAGAGPAPS-----LSGNLQKFILGEILQAPRLLVAH 434 |
| C75278 | A;Accession: C75278 | A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. | Qy 73 PAWHWVTTLSQSAATHH--MAVRDLMRSFLCDMVFYDAMPOQSNLSAFENWATRSA 129 Db 435 PTWGV----DVAEALIHRALDRGATIAVVESELDLFLLSDR-----IAA 480 |
| C75278 | A;Status: preliminary | A;Reference number: A72250; MUID:20036896; PMID:10567266 | Qy 130 LCSPPACSAFPRTISKQDV 149 Db 481 LCSGRUICPAVATASASSQKV 500 |
| C75278 | A;Molecule type: DNA | A;Cross-references: UNIPROT:Q9RR55; GB:AB002071; PMID:96460218; PIDN:AAF1195 | Qy 130 LCSPPACSAFPRTISKQDV 149 Db 481 LCSGRUICPAVATASASSQKV 500 |
| C75278 | A;Genes: DR2410 | A;Map position: 1 | Search completed: June 23, 2005, 08:50:12 |
| C75278 | C:Genetics: | C:Superfamily: DNA-directed DNA polymerase III gamma chain | |

THIS PAGE BLANK (USPTO)

| Scoring table: | BLOSUM22 | | | |
|--|--|--------------------|-------|-------------|
| Gapopen: | 10.0 | | | |
| Gapext: | 0.5 | | | |
| Searched: | 1612378 seqs, 512079187 residues | | | |
| Total number of hits satisfying chosen parameters: | 1612378 | | | |
| Minimum DB seq length: | 0 | | | |
| Maximum DB seq length: | 2000000000 | | | |
| Post-processing: | Minimum Match 0* | | | |
| | Maximum Match 100% | | | |
| | Listing first 45 summaries | | | |
| Database : | UniProt 03 : * 1: uniprot_sprot;*: * 2: uniprot_trembl;*: * | | | |
| Pred. No. | is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. | | | |
| SUMMARIES | | | | |
| Result No. | Score | Query Match Length | DB ID | Description |
| 1 | 2057 | 100.0 | 395 | Q9QUP4 |
| 2 | 1708 | 63.0 | 395 | Q9gxx2 |
| 3 | 1536.5 | 74.7 | 390 | Q7LCN3 |
| 4 | 1536.5 | 74.7 | 411 | Q9GZS9 |
| 5 | 1187 | 57.7 | 392 | Q6GN39 |
| 6 | 1022.5 | 49.7 | 386 | Q8WNCG5 |
| 7 | 1017.5 | 49.5 | 386 | Q9Y5R3 |
| 8 | 1010.5 | 49.1 | 370 | Q8IV46 |
| 9 | 999 | 48.2 | 388 | Q9RI11 |
| 10 | 984 | 47.8 | 388 | Q9WUE5 |
| 11 | 640 | 31.1 | 483 | Q94G9 |
| 12 | 640 | 31.1 | 530 | Q8B276 |
| 13 | 637 | 31.0 | 530 | Q80WV3 |
| 14 | 636.5 | 30.9 | 483 | Q9UED5 |
| 15 | 636.5 | 30.9 | 530 | Q9Y4C5 |
| 16 | 585 | 28.4 | 420 | Q6DBY9 |
| 17 | 581 | 28.2 | 484 | Q8p78 |
| 18 | 580 | 28.2 | 486 | Q75667 |
| 19 | 579 | 28.1 | 486 | Q9NS84 |
| 20 | 578 | 28.1 | 484 | Q9NB0 |
| 21 | 576 | 28.0 | 485 | Q6XQG8 |
| 22 | 563.5 | 27.4 | 479 | Q7LGCS8 |
| 23 | 561.5 | 27.3 | 479 | Q75099 |
| 24 | 551 | 26.8 | 411 | Q9EQCO |
| 25 | 550 | 26.7 | 472 | Q88199 |
| 26 | 549 | 26.7 | 411 | Q43916 |
| 27 | 542 | 26.3 | 474 | Q9QZL2 |
| 28 | 495 | 24.1 | 458 | Q8TCHICK |
| 29 | 457.5 | 22.2 | 441 | Q79403 |
| 30 | 373.5 | 18.2 | 257 | Q79413 |
| 31 | 320.5 | 15.6 | 225 | Q8RY62 |

| | | | |
|----------|---|--|---|
| RP | SEQUENCE FROM N.A. | DR | GO:0008146; F:sulfotransferase activity; IEA. |
| RC | TISSUE=Intestine; | DR | GO:0016740; P:transferase activity; IEA. |
| RA | Submitted (AUG-1998) to the EMBL/GenBank/DBBJ databases. | DR | InterPro: IPR00063; Sulfotransferase_1; 1. |
| RL | EMBL: AF176839; AA056001..1.. | KW | Sulfotransferase. |
| DR | DR | PFam: PF00685; Sulfotransferase_1; 1. | |
| DR | GO; GO:0008146; F:sulfotransferase activity; IEA. | SEQUENCE | 411 AA; 46160 MW; 976422D5BE926E06 CRC64; |
| DR | GO; GO:0016740; P:transferase activity; IEA. | Query Match | 74.7%; Score 1536.5; DB 2; Length 411; |
| DR | DR | Best Local Similarity | 76.0%; Pred. No. 1e-123; |
| DR | InterPro: IPR000867; Sulfotransferase. | Mismatches | 30; Indels 1; Gaps 1; |
| DR | Pfam: PF00685; Sulfotransferase_1; 1. | Matches | 295; Conservative 1; |
| KW | SEQUENCE 390 AA; 43832 MW; 132D1C3B6BE18179 CRC64; | Qy | 1 MRLPFRSSVMLSLMVGQTG-ILYFLVSRQVPSSPAIGLGERHVLVLSWRSGSSFVGQL 59 |
| SQ | Matches 295; Conservative 30; Mismatches 62; Indels 1; Gaps 1; | Db | 60 FSQLPDPVFLMEPAWHWTDLSCGSAPALHMTDLSVFLCDMDVFDAYLPQRNITSD 119 |
| Qy | 1 MRLPFRSSVMLSLMVGQTG-ILYFLVSRQVPSSPAIGLGERHVLVLSWRSGSSFVGQL 59 | Db | 62 FSQPQDPVFLMEPAWHWTDLSCGSAPALHMTDLSVFLCDMDVFDAYPMOSRNLSA 141 |
| Db | 1 MWLPRFSSKTVTTLAQTCTLLPIISRPGPSSPAIGLGERHVLVLSWRSGSSFVGQL 60 | Qy | 120 LFQWAVSRALCSPPVCEAFARGNISSEBVCKPLCATRPEGLAQACEASSYSHVVLKEVRFF 179 |
| Qy | 60 FSQHPDVFYLMEPAMPHWDTLSCGSAPALHMTDLSVFLCDMDVFDAYLPQRNITSD 119 | Db | 142 FFNWATSRALCSPPACASPPRGTSIKQDVCKTLCTROPSSLAREACRSYSHVVLKEVRFF 201 |
| Db | 61 FSQHPDVFYLMEPAMPHWDTLSCGSAPALHMTDLSVFLCDMDVFDAYPMOSRNLSA 120 | Qy | 180 NLQVLYPLISDPAINLRIVHLVEDPRAVLSREOTAKALARDNGIVLGNTGTVWEADPRL 239 |
| Qy | 120 LFQWAVSRALCSPPVCEAFARGNISSEBVCKPLCATRPEGLAQACEASSYSHVVLKEVRFF 179 | Db | 202 NLQVLYPLISDPAINLRIVHLVEDPRAVLSREAGPILARDNGIVLGNTGKWEADPRL 261 |
| Db | 121 FFNWATSRALCSPPACASPPRGTSIKQDVCKTLCTROPSSLAREACRSYSHVVLKEVRFF 180 | Qy | 240 RVVNEVCRESHVRAEAALHKPPPFQLQDYRLVRYEDLARDPLTVIRELYAFTGJGLTPQL 299 |
| Qy | 180 NLQVLYPLISDPAINLRIVHLVEDPRAVLSREOTAKALARDNGIVLGNTGTVWEADPRL 239 | Db | 262 RLIREVCRESHVRAEAATLKKPPPLRGRVLRVEDPRAEIRALYTAFTGJGLTPQL 321 |
| Db | 181 NLQVLYPLISDPAINLRIVHLVEDPRAVLSREAGPILARDNGIVLGNTGKWEADPRL 240 | Qy | 300 QTWHNITHSGPGCARERAFKITSRDAALSQWHTLPPAKIRRVQELGGALQLGVR 359 |
| Qy | 240 RVVNEVCRESHVRAEAALHKPPPFQLQDYRLVRYEDLARDPLTVIRELYAFTGJGLTPQL 299 | Db | 322 EAIIHNTITHSGGKPKIEFHFTSSRNARNVSQAWRHALPFTKLRLVQEYCAGALQLGVR 381 |
| Db | 241 RLIREVCRESHVRAEAATLKKPPPLRGRVLRVEDPRAEIRALYTAFTGJGLTPQL 300 | Qy | 360 SVHSEBLEQNDLSLDLPPGMDSPKWAS 387 |
| Qy | 300 QTWHNITHSGPGCARERAFKITSRDAALSQWHTLPPAKIRRVQELGGALQLGVR 359 | Db | 382 PVYSAQQDQLTDLVLPKGPDFHSWAS 409 |
| Db | 301 EAIIHNTITHSGGKPKIEFHFTSSRNARNVSQAWRHALPFTKLRLVQEYCAGALQLGVR 360 | RESULT 5 | |
| Qy | 360 SVHSEBLEQNDLSLDLPPGMDSPKWAS 387 | Q6GN39 | PRELIMINARY; |
| Db | 361 PVYSAQQDQLTDLVLPKGPDFHSWAS 388 | ID | PRT; 392 AA. |
| RESULT 4 | | | |
| 09G259 | Q9GZ59 PRELIMINARY; | AC | Q6GN39; 05-JUL-2004 (TREMBLrel. 27, Created) |
| ID | AC | DT | 05-JUL-2004 (TREMBLrel. 27, Last sequence update) |
| AC | Q9GZ59; Q9UBY3; | DE | MGCB3048 Protein. |
| DT | 01-MAR-2001 (TREMBLrel. 16, Created) | GN | Name=MGCB3048; |
| DT | 01-MAR-2001 (TREMBLrel. 16, Last sequence update) | OS | Xenopus laevis (African clawed frog). |
| DT | 05-JUL-2004 (TREMBLrel. 27, Last annotation update) | OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; |
| DE | [1] _TaxID=8355; | OC | Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Xenopus. |
| DE | NCBI_TaxID=8355; | OX | NCBI_TaxID=8355; |
| DE | RN [1] | RN | SEQUENCE FROM N.A. |
| DE | RP SEQUENCE FROM N.A. | RC TISSUE=Embryo; | |
| OS | Homino sapiens (Human); Chordata; Craniata; Vertebrata; Buteleostomi; | RX MEDLINE=22383257; PubMed=12477932; DOI=10.1073/pnas.242603899; | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; | RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., | |
| OC | Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. | RA Klausner R.D., Collins F.S., Wagner L., Shanmen C.M., Schuler G.D., | |
| OX | [1] _TaxID=9606; | RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., | |
| RN | SEQUENCE FROM N.A. | RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hosieh F., | |
| RP | MEDLINE=20472310; PubMed=11017086; DOI=10.1038/79987; | RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., | |
| RX | AKama T.O., Nishida K., Nakayama J., Watanabe H., Ozaki K., Nakamura T., Dots A., Maeda N., Yamamoto S., Fujiiwara T., Thonar E.J., Shimomura Y., Kinoshita S., Tanigami A., Fukuda M.N., | RA Stapleton M., Soares M.B., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Rahja S.S., Logueilano N.A., Peters G.J., Abramson R.D., Mullah S.J., Boesk S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muny D.M., Sodergren B.J., Lu X., Gibbs R.A., Fahey J., Heitton E., Ketteman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green B.D., Dickason M.C., Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S., | |
| RL | Nat. Genet. 26:237-241(2000). | RA | |
| DR | EMBL: AF246718; AAC28023..1.. | RA | |
| DR | EMBL: AF219991; AAC26326..1.. | RA | |
| DR | Genev; HGNC:1973; CHST5. | RA | |

| | |
|--------|--|
| RA | Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K., Masuho Y., Ono T., Okano K., Yoshihara Y., Aotsuka S., Sasaki N., Hattori A., Okumura K., Iwayangji T., Ninomiya K.; |
| RA | Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases. |
| RL | EMBL; AK07446 ; BAC11177.1 ; |
| DR | GeneID: HGNC:1972 ; CHSP4. |
| DR | GO: 0008146; P:sulfotransferase activity; IFA. |
| DR | GO: 1016740; P:transferase activity; IFA. |
| DR | InterPro: IPR000663; Sulfotransferase. |
| Pfam | PF00655; Sulfotransfer_1. |
| KW | Transf erase. |
| SQ | SEQUENCE 386 AA: 45160 MW: 861869348319E42A CRC64; |
| | Query Match 49.7%; Score 1022.5; DB 2; Length 386; |
| | Best Local Similarity 56.7%; Prod. No. 1.1e-79; |
| | Matches 216; Conservative 42; Mismatches 112; Indels 8 11; Gaps 7; |
| Qy | 1 MRLPFRSSVMLSLMVGOTGILVF---LVSROVPS-SPAGLGERVHVLVLSRSRSGSSPV 56 |
| Db | 1 MLIPK--KMLPKLFLVSSOMAILAFFMYSHNISSSLMSKAQPERMHVLVLSRSRSGSSPV 58 |
| Qy | 57 GOLFSQHPDVYFLMEPAWHWDLTSQGSAPALIMAVRDLLIRSLVFLCDMDVFDAYL-PWRR 115 |
| Db | 59 GQFLGQHDFDVYFLMEPAWHWDLTSQGSAPALIMAVRDLLIRSLVFLCDMDVFDAYLMEPGPR 118 |
| Qy | 116 NISDLFQWAVSRLCSPVCEAFARGNISSEEVCKPLCATRPGGLAEQACSSYSHVVLKE 175 |
| Db | 119 RQSSLFQWENSRALCSAACDLPQDELETPRAHCRLCSQQPEEVEXACRS1SHVVLKE 178 |
| Qy | 176 VRFENLQVYLPLISDPA1NLRYHLVLDRAVRSREQTAKALARNDIVLGTINGTW-B 234 |
| Db | 179 VRFENLQSLQSLPLKDPSENLIYHLVLDRAVRSREITKGDLMDNR1VMGHEQKLKK 238 |
| Qy | 235 ADPRLRVNEYCRSHVRTEAAALHKPPFLQDORYLRYEDLARDPLTVIRELYAFTGIG 294 |
| Db | 239 EDQPPYYQVQIVCOSQLEYK-TIQSLPKALQBEQYLLRYEDLARAPQAQTQRSMYBFGVJL 297 |
| Qy | 295 LTPQLQTWHNITHGSGPGRARFAKFTTSRDLASVQSWRHTLPPAKIIRRQVQELCGGALQ 354 |
| Db | 298 FLPHLCQTWHNITRGKMG--DHAFTNARDALNVQSWARWSLPEYKRSRLOQAGDMN 355 |
| Qy | 355 LIGYRSTVSELEBORDLSIDL 375 |
| Db | 356 LIGYRHTVSEQEORNLIDL 376 |
| | RESULT 7 |
| QYYSR3 | PRELIMINARY; |
| ID | Q9Y5R3 |
| AC | Q9Y5R3; |
| DT | 01-NOV-1999 (TREMBLrel. 12, Created) |
| DT | 01-NOV-1999 (TREMBLrel. 12, Last sequence update) |
| DT | 05-JUL-2004 (TREMBLrel. 27, Last annotation update) |
| DE | N-acetylglucosamine 6-O-sulfotransferase (L-selectin ligand sulfotransferase GST-3). |
| DE | Name=GST3; |
| GN | Homo sapiens (Human) |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. |
| OX | NCBI_TaxID=9606; |
| RN | SEQUENCE FROM N_A. |
| RC | SEQUENCES:1; |
| RD | Medline# 99264336; PubMed=10330415; DOI=10.1083/jcb.145.4.899; |
| RA | Bistrip A., Phak S., Lee J.U., Belov Y.Y., Gunn M.D., Zuo F.R., Huang C.C., Kanagi R., Rosen S.D., Hemmerich S.; |
| RT | "Sulfotransferases of two specificities function in the reconstitution of high endothelial cell ligands for L-selectin."; |
| RL | J. Cell. Biol. 145:899-910(1999). |
| RN | SEQUENCE FROM N_A. |
| RC | SEQUENCES:1; |

| | | | |
|---------------------------|--|--------------------|--|
| GN | Mus musculus (Mouse) | SEQUENCE FROM N.A. | 242 YYAMKICKSQDVDTIVK-AIQTLPEALQQQYFLRVDFLPLPH 300 |
| OS | Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus. | | |
| OC | NCBI_TaxID=10090; | | |
| OX | [1] | | |
| RN | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=93161934; PubMed=10435581; DOI=10.1016/S1074-7613(00)80083-7; | | |
| RA | Hirakawa N., Petrynia B., Nakayama J., Tsutsumi S., Suzuki M., Yeh J.C., Iwaza D., Tanaka T., Miyazaki M., Lowe J.B., Fukuda M.; | | |
| RA | "A novel, high endosomal membrane-specific sulfotransferase expresses 6-sulfotransfase (Lewis IX), an L-selectin ligand displayed by CD34."; | | |
| RT | RT | RESULT 10 | |
| RL | Immunity 11:79-89(1999). | | |
| RN | SEQUENCE FROM N.A. | | |
| RP | STRAIN=FVB/N; TISSUE=Colon; | | |
| RC | MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; | | |
| RX | Strain R.L., Feingold B.A., Grouse L.H., Derge J.G., Klauner R.D., Collins F.S., Wagner L., Shenmen C.M., Schulter G.D., Altenschul S.F., Zeeberg B., Buerow K.H., Schaeffer C.F., Bhat N.K., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Tsohomyan C., Carminci P., Prange C., Loqueland N.A., Peters G.J., Abramson R.D., Mulitaly S.J., Raha S.S., McEvian P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hile S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting M., Maden A., Kettman M., Maden A., Rodriguez S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimes W.J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smialius D.E., Schnierch A., Schein J.R., Jones S.J., Marras M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). | | |
| RN | SEQUENCE FROM N.A. | | |
| RP | STRAIN=FVB/N; TISSUE=Colon; | | |
| RC | MEDLINE=20031; Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases. | | |
| RA | Strausberg R.; Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases. | | |
| DR | EMBL: AF101155; AAD45579.1; -; | | |
| DR | EMBL: BC057886; AAH57886.1; -; | | |
| DR | MGD: MGI_139479; Chat4. | | |
| DR | GO: GO:0016021; C: integral to membrane; TAS. | | |
| DR | InterPro: IPR000863; Sulfotransferase. | | |
| DR | Pfam: PF00855; Sulfotransferase_1. | | |
| KW | Selectin; Transferase. | | |
| SEQUENCE | 388 AA; 44635 MW; 6D5371AFB6884ABE CRC64; | | |
| Query Match | 48.2%; Score 991; DB 2; Length 388; | | |
| Best Local Similarity | 55.2%; Pred. No. 9.1e-77; | | |
| Matches 208; Conservative | 43; Mismatches 114; Indels 12; Gaps 6; | | |
| Qy | 2 RLPFR-SSTVMLSLLMVTGILYFLVSROPVSSPAGLERVAVYLSSWRGSSSFVQGF 60 | | |
| Db | 8 RALMFGLSQVFSVTHRSQREESRRP-----VTVLVLSSWRGSSSFVQGF 61 | | |
| Qy | 61 SQRDPVFLMEPAWHWDTLSQSOPASPAHMYARDLISVFLCDMDVDAVL-FWRRNSD 119 | | |
| Db | 62 GGHPDVFLMEPAWHWMTTSSTAKHMAVDLLESVFLCDMSVDAVMNPGRKQSS 121 | | |
| Qy | 120 IFOQWA9SRLASCPSPVCEAFARGNISSEEVCKPLCATRPFGLAOBACSSYSHVVLKEVRFF 179 | | |
| Db | 122 IFQWEQSRALLASCPVCDFFPAHEISSPAHKCLIGQQPFDWVACRSHGTFLVKEVRFF 181 | | |
| Qy | 180 NIQVLYPLLSDPALNLRIVHLVYDPRAVTRSRREQTAKALDRNGTIVLGTN-STWVEADPR 238 | | |
| Db | 182 SIQALPFLSLLPDSLNHhvHVLVYDPRAVTRSRREQTAKALDRNGTIVLGTN-STWVEADPR 241 | | |
| Qy | 239 IFRVNNFCRSFHRIAEALAHKKPPFLQDRYRURYEDLARDPLTVIRELAFTQLGTPQ 298 | | |

DE N-acetylglucosamine-6-O-sulfotransferase long form.
 OC Name=Chat2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1];
 SEQUENCE FROM N.A.
 RP TISSUE:Whole embryo;
 RX MEDLINE:98380482; PubMed=9712885; DOI=10.1074/jbc.273.35.22577;
 RA Ichimura K., Muramatsu H., Kadomatsu K., Fan Q., Kurokawa N.,
 RA Mitsuoka C., Kannagi R., Habuchi O., Muramatsu T.;
 RT "Molecular Cloning and characterization of an N-acetylglucosamine-6-O-
 sulfotransferase."
 RT J. Biol. Chem. 273:22577-22583 (1998).
 DR EMBL; AB011152; BAA3238.1; -.
 DR MCB; BAA3238.1; Chat2.
 DR GO:0005615; C:extracellular space; TAS.
 DR GO:0008146; P:sulfotransferase activity; IDA.
 DR GO:0006044; P:N-acetylglucosamine metabolism; TAS.
 DR InterPro:IPR000663; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer_1; 1.
 KW Transferase.
 SQ SEQUENCE 530 AA; 57814 MW; A113E1B735C363EC CRC64;

Query Match Score 640; DB 2; Length 530;
 Best Local Similarity 36.2%; Pred. No. 2, 4e-46;
 Matches 145; Conservative 73; Mismatches 118; Indels 64; Gaps 10;

Qy 30 VPSSPAG-----LG-----BRVHVLVLSMRSGSSFVGOLFQSHQPDVFYLFME 71
 Db 136 VPAAAAGSAGAAASLGNATRGTRGGDKRQLVYFTWRSGSSFFGELFNQNPEVFFLYE 195

Qy 72 PATHWDTLSQGASAPALHMAVDRDLIRSVFLCDMDVFDAYLP--WRNISDL--FOWAVS 126
 Db 196 PWVHWWORKLPGDAVSLQGARDMLSLARYCDLSVFPOLYSPACSGGRNLTTGIGATN 255

Qy 127 RALCSPPVCEAFARG--NISSEEVCKPLCATRPFGLAQEACSSYSHVTLKEVRFNQLV 184
 Db 256 KVCCSSPLCPAYKEVVLGUDRVCKK--CPQPQLARFEECKRTVVKGYRVFDVAL 314

Qy 185 YPLISDPAINLRIVHLVRDPRAYLRSREQATAKALARDNGIVLTGTNGTWADPR---- 238
 Db 315 APLKDPDALKTVHLVRDPRAVASSRIRSRLQVRSR----DERAHRMPF 368

Qy 239 -----LRYVNECRSHVRLAEALHKPPFFLQDQRYLRYED 275
 Db 369 LEAAGHKUGAKCEGMGPADDHALGAMEVCINSMAKTLQTAL.QPPDWLQGHLYVRYED 427

Qy 276 LARDPLTVIRELYAFTGCLPTQLWTIHNSGPARAFKTSRDALSVSQWRH 335
 Db 428 LVGPVKTIRYDFVGLVSPMEMEOFALNMSGSSSK--PFVVSARNATQAANAWRT 485

Qy 336 TLPPAKIRRQEILGGALOLGYRSVHSELORDLSLDLL 375
 Db 486 ALTFQQIKQVEEFCYQPMALVGYERVSPEVKDSLSTLL 525

RESULT 13
 Q80WV3 PRELIMINARY; PRT; 530 AA.
 ID Q80WV3;
 AC Name=Chat2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-MAR-2004 (TrEMBLrel. 24, Last sequence update)
 DT Chat2_Protein.
 GN Name=Chat2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1];
 RP SEQUENCE FROM N.A.

RESULT 14
 Q9UED5 PRELIMINARY;
 ID Q9UED5;
 AC Name=Chat2;

RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=2238825; PubMed=1077932; DOI=10.1073/pnas.24203899;
 RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.; Gruisken C.M.; Schuler G.D., Klausner R.D.; Collins P.S.; Wagner L.; Shemesh T.; Mullahy S.J.; Altshul S.F.; Zeeberg B.; Buetow K.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.; Hopkins R.F.; Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.; Stapleton M.; Bonaldo M.F.; Casarotto T.L.; Scheetz T.E.; Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.; Raha S.S.; Loqueland N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.; Bosak S.A.; McEvany P.J.; McKernan R.D.; Gunaratne P.H.; Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Gibbs R.A.; Villa-On D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.; Fahey J.; Helton E.; Kettman M.; Madan A.; Rodriguez S.; Sanchez A.; Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.; Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.; Mullahy S.J.; McEvany P.J.; Gunaratne P.H.; Mullahy S.J.; Chat2.
 RA Rodriguez A.C.; McEvany P.J.; Gunaratne P.H.; Mullahy S.J.; Chat2.
 RA Krzywinski M.I.; Skalska U.; Smailus D.E.; Schnurch A.; Jones S.J.; Marra M.A.; RT "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [12];
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.; Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases;
 RL EMBL; BC051963; AAH51963.1; -.
 DR MGD; MGI:1891167; Chat2.
 DR GO:0005615; C:extracellular space; TAS.
 DR CO: GO:0008146; P:sulfotransferase activity; IDA.
 DR InterPro: IPR000863; Sulftotransferase.
 DR Pfam: PF00685; Sulfotransfer_1; 1.
 SEQUENCE 530 AA; 57828 MW; 27533BF15440730 CRC64;
 SEQUENCE 530 AA; 57828 MW; 27533BF15440730 CRC64;
 Query Match Score 637; DB 2; Length 530;
 Best Local Similarity 36.0%; Pred. No. 4.3e-46;
 Matches 144; Conservative 74; Mismatches 118; Indels 64; Gaps 10;
 DR 30 VPSSPAG-----LG-----BRVHVLVLSMRSGSSFVGOLFQSHQPDVFYLFME 71
 DR 136 VPAAAAGSAGAAASLGNATRGTRGGDKRQLVYFTWRSGSSFFGELFNQNPEVFFLYE 195
 DR 72 PAWHWDTLSQGASAPALHMAVDRDLIRSVFLCDMDVFDAYLP--WRNISDL--FOWAVS 126
 DR 196 PWVHWWORKLPGDAVSLQGARDMLSLARYCDLSVFPOLYSPACSGGRNLTTGIGATN 255
 DR 127 RALCSPPVCEAFARG--NISSEEVCKPLCATRPFGLAQEACSSYSHVTLKEVRFNQLV 184
 DR 256 KVCCSSPLCPAYKEVVLGUDRVCKK--CPQPQLARFEECKRTVVKGYRVFDVAL 314
 DR 185 YPLISDPAINLRIVHLVRDPRAYLRSREQATAKALARDNGIVLTGTNGTWADPR---- 238
 DR 315 APLKDPDALKTVHLVRDPRAVASSRIRSRLQVRSR----DERAHRMPF 368
 DR 239 -----LRYVNECRSHVRLAEALHKPPFFLQDQRYLRYED 275
 DR 369 LEAAGHKUGAKCEGMGPADDHALGAMEVCINSMAKTLQTAL.QPPDWLQGHLYVRYED 427
 DR 428 LVGPVKTIRYDFVGLVSPMEMEOFALNMSGSSSK--PFVVSARNATQAANAWRT 485
 DR 486 ALTFQQIKQVEEFCYQPMALVGYERVSPEVKDSLSTLL 525
 DR 276 LARDPLTVIRELYAFTGCLPTQLWTIHNSGPARAFKTSRDALSVSQWRH 335
 DR 428 LVGPVKTIRYDFVGLVSPMEMEOFALNMSGSSSK--PFVVSARNATQAANAWRT 485
 DR 315 APLKDPDALKTVHLVRDPRAVASSRIRSRLQVRSR----DERAHRMPF 368
 DR 239 -----LRYVNECRSHVRLAEALHKPPFFLQDQRYLRYED 275
 DR 369 LEAAGHKUGAKCEGMGPADDHALGAMEVCINSMAKTLQTAL.QPPDWLQGHLYVRYED 427
 DR 428 LVGPVKTIRYDFVGLVSPMEMEOFALNMSGSSSK--PFVVSARNATQAANAWRT 485
 DR 336 TLPPAKIRRQEILGGALOLGYRSVHSELORDLSLDLL 375
 DR 486 ALTFQQIKQVEEFCYQPMALVGYERVSPEVKDSLSTLL 525

| | | |
|--|---|---|
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) | RESULT 15 |
| DT | 05-JUL-2004 (TREMBLrel. 23, Last sequence update) | Q9Y4C5 PRELIMINARY; PRT; 530 AA. |
| DE | N-acetylglucosamine-6-O-sulfotransferase (GlcNAc6ST). | ID Q9Y4C5; Q9GZN5; 09Y672; |
| GN | Name=Gn6ST; | AC DT 01-NOV-1993 (TREMBLrel. 12, Created) |
| OS | Homo sapiens (Human). | DT DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update) |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | DT DT 05-JUN-2004 (TREMBLrel. 27, Last annotation update) |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | DB Long form of N-acetylglucosamine-6-O-sulfotransferase (GlcNAc6ST) |
| OX | NCBI_TaxID=9606; | DE (Carbohydrate sulfotransferase 2). |
| RN | | GN Name=Gn6ST; Synonyms=CHST2; |
| RP | SEQUENCE FROM N.A. | OS Homo sapiens (Human); |
| RC | TISSUE:Brain; | OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| RX | Medline:98391845; PubMed=9722682; | GN TAXID=9606; |
| RA | Ochiai K., Muramatsu H., Yamakawa T., Fan Q., | RN |
| RA | Uchimura K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q., | RP SEQUENCE FROM N.A. |
| RA | Ozaki T., Nakagawa A., Kadomatsu K., Muramatsu T., | RC TISSUE:Brain; |
| RA | "Human N-acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of 6-sulfat sialyl Lewis X: molecular cloning, chromosomal mapping, and expression in various organs and tumor cells."; | MEIDLINE=98391845; PubMed=9722682; |
| RL | Uchimura C., Kannagi R., Habuchi O., Yokoyama I., Yamamura K., | RA Uchimura K., Kannagi R., Yokoyama I., Yamamura K., Muramatsu T., |
| RA | Mitsuoaka C., Nakagawa A., Kadomatsu K., Muramatsu T., | RA Mitsuoka C., Nakagawa A., Kadomatsu K., Muramatsu T., |
| RA | Sakaguchi H., Kitagawa H., Sugahara K.; | RA Ozaki T., Nakagawa A., Kadomatsu K., Muramatsu T., |
| RA | "Functional expression and genomic structure of human N-acetylglucosamine-6-O-sulfotransferase that transfers sulfate to b-N-acetylgalactosamine at the nonreducing end of an N-acetyllactosamine sequence."; | RT "Human N-acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of 6-sulfat sialyl Lewis X: molecular cloning, chromosomal mapping, and expression in various organs and tumor cells."; |
| RL | Biochem. 124:670-678(1998). | RL Biochem. 124:670-678(1998). |
| RP | SEQUENCE FROM N.A. | RN |
| RC | TISSUE:Placenta; | RP SEQUENCE FROM N.A. |
| RX | Medline:90498786; PubMed=11042394; DOI=10.1016/S0304-4165(00)00136-7; | RC TISSUE:Umbilical vein endothelium; |
| RA | Ozaki T., Nakagawa A., Kadomatsu K., Muramatsu T., | MEIDLINE=99168905; PubMed=10049591; DOI=10.1006/geno.1998.5653; |
| RA | "Functional expression and genomic structure of human N-acetylglucosamine-6-O-sulfotransferase that transfers sulfate to b-N-acetylgalactosamine at the nonreducing end of an N-acetyllactosamine sequence."; | RA Li X., Tedder T.F.; |
| RT | RT | RA "CHST1 and CHST2 sulfotransferases expressed by human vascular endothelial cells: cDNA cloning, expression, and chromosomal localization."; |
| RT | RT | RA Genomics 55:345-347(1999). |
| RT | RT | DR AB014680; BAA34266; - |
| RT | RT | DR AF083066; AAD2081.1; - |
| RT | RT | DR Genew; HGNC:1970; CHST2. |
| RL | RL | DR GO; GO:000145; P:sulfotransferase activity; IEA. |
| DR | DR | DR GO; GO:0016740; P:transferase activity; IEA. |
| DR | DR | DR InterPro; IPR00863; Sulfotransfser_1. |
| DR | DR | KW Transferase activity; IEA. |
| KW | | DR Pfam; PF00685; Sulfotransfser_1. |
| SQ | SEQUENCE 483 AA: B27797D4931BA18 CRC64; | SQ SEQUENCE 530 AA: 57857 MW: A82CA227B9D5651B CRC64; |
| Query | Query Match Score 30.9%; Length 483; | Query Match Score 30.9%; Length 530; |
| Best Local Similarity 37.1%; Pred. No. 4.2e-46; | Best Local Similarity 37.1%; Pred. No. 4.7e-46; | Best Local Similarity 37.1%; Pred. No. 4.7e-46; |
| Matches 140; Conservative 73; Mismatches 117; Indels 47; Gaps 9; | Matches 140; Conservative 73; Mismatches 117; Indels 47; Gaps 9; | Matches 140; Conservative 73; Mismatches 117; Indels 47; Gaps 9; |
| Db | Db | Db |
| Qy | 36 GLGE-RVHVVLSSWRSGSSFVGOLFSDQHDPVYLMPEAWHWDTLSCGSAPALHMYRD 94 | Qy 36 GLGE-RVHVVLSSWRSGSSFVGOLFSDQHDPVYLMPEAWHWDTLSCGSAPALHMYRD 94 |
| Qy | 112 GVGDRQLRVFTFWRSGSSFFGAGSGGNNLTIGATGATKVKVCSPLCPAYKEYVGLVDR 171 | Qy 112 GVGDRQLRVFTFWRSGSSFFGAGSGGNNLTIGATGATKVKVCSPLCPAYKEYVGLVDR 171 |
| Db | Db | Db |
| Qy | 95 LIRSYVLCDMDVFDAYLP--WRRNISDL--FQAVASRALSCLSPVCEAFARG--NISSEE 147 | Qy 95 LIRSYVLCDMDVFDAYLP--WRRNISDL--FQAVASRALSCLSPVCEAFARG--NISSEE 147 |
| Db | 172 MLSALYRCDSLPSQPLSPAGSGGNNLTIGATGATKVKVCSPLCPAYKEYVGLVDR 231 | Db 172 MLSALYRCDSLPSQPLSPAGSGGNNLTIGATGATKVKVCSPLCPAYKEYVGLVDR 231 |
| Qy | 148 VCKPLCATREFGLAQEACSSSHVVLKBYFRFNIVQVLYPLSLSPALNIRVHLVRDPRAV 207 | Qy 148 VCKPLCATREFGLAQEACSSSHVVLKBYFRFNIVQVLYPLSLSPALNIRVHLVRDPRAV 207 |
| Db | 232 VCKKL-CPPQBLARFEECRKYRTLVIGTRVEDAVALLRLRDPALDKVHLVRDPRAV 290 | Db 232 VCKKL-CPPQBLARFEECRKYRTLVIGTRVEDAVALLRLRDPALDKVHLVRDPRAV 290 |
| Qy | 208 LRSREQTAKLARDNGTGTWEDPR-----238 | Qy 208 LRSREQTAKLARDNGTGTWEDPR-----238 |
| Db | 291 ASSRIRSRLIRESLQVRSR-----DRAHMPFLAAGHKLGAKKEGVGGPADYHA 344 | Db 291 ASSRIRSRLIRESLQVRSR-----DRAHMPFLAAGHKLGAKKEGVGGPADYHA 344 |
| Qy | 239 LRVNECRSHVRIAEEALKKPPPLQDRXRLVRYEDLARDPLTVIRELYAFGLGLTPO 298 | Qy 239 LRVNECRSHVRIAEEALKKPPPLQDRXRLVRYEDLARDPLTVIRELYAFGLGLTPO 298 |
| Db | 345 LGAMEVICNSMALKLQLTAQPPDMLQHLYVRYEDLQGDPUVTLRRLVYDFVGLVSP 403 | Db 345 LGAMEVICNSMALKLQLTAQPPDMLQHLYVRYEDLQGDPUVTLRRLVYDFVGLVSP 403 |
| Qy | 299 LQTWHNITHGSGPGRARREFKTSRDAVSQAWRHTLEFAKIRRVQBLGGALQLGY 358 | Qy 299 LQTWHNITHGSGPGRARREFKTSRDAVSQAWRHTLEFAKIRRVQBLGGALQLGY 358 |
| Db | 404 MEQFALNTSGS6SSSK--FVVVSARNATQAANAWRTALTFOQIKQVEFBYCQPMALVGY 461 | Db 404 MEQFALNTSGS6SSSK--FVVVSARNATQAANAWRTALTFOQIKQVEFBYCQPMALVGY 461 |
| Qy | 359 RSVHSELEQDLSLDL 375 | Qy 359 RSVHSELEQDLSLDL 375 |
| Db | 462 ERVNSPEEVQDLSKTL 478 | Db 462 ERVNSPEEVQDLSKTL 478 |
| Qy | 359 RSVHSELEQDLSLDL 375 | Qy 359 RSVHSELEQDLSLDL 375 |

Db 509 | :| :||| |||
509 ERVNSPEEVRLSKTLL 525

Search completed: June 23, 2005, 08:48:57
Job time : 46.3472 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein search, using sw model

Run on: June 23, 2005, 08:32:43 ; Search time 42.6885 Seconds

(without alignments)
3578.727 Million cell updates/sec

Title: US-10-697-828-7

Perfect score: 2057

Sequence: 1 MRLPRSSVTMISLMMQTS.....LPRGMDSFKMASSTEKQPES 395

Scoring table: BL0SUM62

Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: GeneseqD1980s:*
- 2: GeneseqD1990s:*
- 3: GeneseqD2000s:*
- 4: GeneseqD2001s:*
- 5: GeneseqD2002s:*
- 6: GeneseqD2003aa:*
- 7: GeneseqD2003ba:*
- 8: GeneseqD2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Match Length | DB ID | Description |
|------------|--------|-------|--------------|------------|--------------------|
| 1 | 2057 | 100.0 | 395 | 4 PAV72638 | Aay72638 Mouse gly |
| 2 | 2057 | 100.0 | 395 | 5 AAU1175 | Aau1175 Murine int |
| 3 | 2042.5 | 99.3 | 418 | 5 ABB81557 | Abb81557 Mouse int |
| 4 | 17227 | 84.0 | 395 | 5 ABB81555 | Abb81555 Consensus |
| 5 | 1708 | 83.0 | 395 | 4 AAY72640 | Aay72640 Human gly |
| 6 | 1708 | 83.0 | 395 | 5 ABB81554 | Abb81554 Human cor |
| 7 | 1708 | 83.0 | 395 | 5 AAE15438 | Aae15438 Human dru |
| 8 | 1708 | 83.0 | 395 | 7 ADI21086 | Adi21086 Novel hum |
| 9 | 1708 | 83.0 | 395 | 8 ADL61235 | Adl61235 Human tyr |
| 10 | 1536.5 | 74.7 | 390 | 4 PAV72639 | Aay72639 Human gly |
| 11 | 1536.5 | 74.7 | 390 | 5 ABB81556 | Abb81556 Human int |
| 12 | 1436 | 69.8 | 418 | 3 AAB41947 | Aab41947 Human ORF |
| 13 | 1333 | 64.8 | 394 | 7 ADJ70405 | Adj70405 Human hea |
| 14 | 1022.5 | 49.7 | 386 | 4 AAM93309 | Aam93309 Human pol |
| 15 | 1022.5 | 49.7 | 386 | 8 ADL30784 | Adl30784 Human pro |
| 16 | 1017.5 | 49.5 | 386 | 2 AAY39918 | Aay39918 Human gly |
| 17 | 986.5 | 49.4 | 380 | 5 AAU11274 | Aau11274 Human L-s |
| 18 | 986.5 | 48.0 | 386 | 3 AAY79219 | Aay79219 Human tra |
| 19 | 984 | 47.8 | 388 | 2 AAY39919 | Aay39919 Mouse gly |
| 20 | 712 | 34.6 | 169 | 5 ABB81558 | Abb81558 Human cor |
| 21 | 677 | 32.9 | 169 | 5 ABB81559 | Abb81559 Human int |
| 22 | 647.5 | 31.5 | 484 | 2 AAY31657 | Aay31657 Human N-a |
| 23 | 647.5 | 31.5 | 531 | 4 AAU6914 | Aau6914 Lung smal |
| 24 | 647.5 | 31.5 | 531 | 8 ADR14283 | Adr14283 Human NF- |
| 25 | 640 | 31.1 | 483 | 2 AAY31656 | Aay31656 Mouse N-a |

ALIGNMENTS

| | | | | | |
|----|-------|------|-----|-------------|---------------------|
| 26 | 636.5 | 30.9 | 530 | 4 AAB95367 | Ab95367 Human pro |
| 27 | 636.5 | 30.9 | 530 | 8 ADQ18590 | Adq18590 Human sof |
| | | | | 6 ABP56121 | Abp56121 Human cho |
| | | | | 6 ABP5237 | Abp5237 Human pro |
| | | | | 6 ABR41139 | Abr41139 Human DTR |
| | | | | 6 AAW52863 | Aaw52863 Glycosami |
| | | | | 2 ABB81556 | Abb81556 Mouse pro |
| | | | | 7 ABB81552 | Abb81552 Human BEC |
| | | | | 7 ADN95522 | Adn95522 Human |
| | | | | 7 AAW61100 | AAw61100 Keratan s |
| | | | | 5 AAc25356 | Aac25356 Human cho |
| | | | | 5 ABU03503 | Abu03503 Angiogene |
| | | | | 6 ADJ61589 | Adj61589 Human hea |
| | | | | 7 ADJ618725 | Adj618725 Human sof |
| | | | | 8 ABB81560 | Abb81560 Human hig |
| | | | | 7 ADI21571 | Adi21571 Novel hum |
| | | | | 5 AAB80522 | Aab80522 Human dia |
| | | | | 8 AAW06480 | Aaw06480 Chick cho |
| | | | | 5 ABBB1561 | AbbB1561 Human N-a |
| | | | | 5 ABBB1563 | AbbB1563 Human cho |
| | | | | 5 ABBB1562 | AbbB1562 Human ker |

RESULT 1

AAV72638 standard; protein: 395 AA.

ID AAV72638;

XX AC;

XX DT;

02-MAY-2001 (first entry)

XX DE;

XX KW;

KW selectin binding inhibitor; gene therapy; inflammation;

KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;

KW polyarthritis nodosa; polymyositis; systemic sclerosis; dermatitis;

KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; Hashimoto's disease; Graves' disease; hypoparathyroidism; anaemia; denyeinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;

KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;

KW asthma; hypersensitivity; rheumatic fever; tissue rejection;

KW chromosome 8E1.

XX OS;

OS Mus musculus.

XX PN;

PN W0200106015-A1.

XX PD;

PD 25-JAN-2001.

XX PP;

PP 19-JUL-2000; 2000WO-US019741.

XX XX;

XX PR 20-JUL-1999; 99US-0144694P.

PR 13-JUN-2000; 2000US-00593828.

XX PR WPI; 2001-138471/14.

DR N-SDB; AAD02696.

XX PI; Rosen SD, Lee JK, Hemmerich S;

XX XX;

XX DR; UNIV CALIFORNIA.

XX PA;

PA New Glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications.

XX PS; Claim 3; Fig 2; 128pp; English.

CC The present sequence is mouse glycosyl sulfotransferase-4 (GST-4). GST-4 gene is found on chromosome 8E1. GST is a type 2 membrane protein, useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises a contacting the selectin with a non-sulphated selectin

CC

CC

CC

CC

CC

CC

CC ligand, GST and a small molecular agent that inhibits the sulphation
 CC activity of GST. GST is also useful in inhibiting a selectin mediated
 CC binding event. GST is useful in gene therapy to treat disorders such as
 CC acute or chronic inflammation, systemic lupus erythematosus (SLE),
 CC rheumatoid arthritis, polyarteritis nodosa, polymyositis,
 CC dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis,
 CC myasthenia gravis, Sjögren's syndrome, Hashimoto's disease, Grave's
 CC disease, adrenalitis, hypoparathyroidism, pernicious anaemia,
 CC demyelinating diseases, cirrhosis, ulcerative colitis, dermatitis,
 CC myocarditis, regional enteritis, adult respiratory distress syndrome,
 CC infantile eczema, psoriasis lichen planus, allergic rhinitis, bronchial
 CC asthma, hypersensitivity, rheumatic fever and tissue rejection during
 XX transplantation
 CC Sequence 395 AA:

```
Query Match          Score 2057; DB 4; Length 395;
Best Local Similarity      100.0%; Pred. No. 1.7e-205;
Matches      395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy          1 MRLPRESSTVMSLIMVGTQSLIVELVSQVPSSPAGLGERVHVLVLSWRSGSSFVGOLF 60
Db          1 MRLPRESSTVMSLIMVGTQSLIVELVSQVPSSPAGLGERVHVLVLSWRSGSSFVGOLF 60
Qy          61 SQHDPVYFLMEPAWHDVTLISQSSAPALHMAYDLDLRSVFLCDMDVFDAYLPWRNISDL 120
Db          61 SQHDPVYFLMEPAWHDVTLISQSSAPALHMAYDLDLRSVFLCDMDVFDAYLPWRNISDL 120
Qy          121 FQAVWSRALCSPPVCEAFARGNISSEEVCKPLCATEPRPGLAQACSSYSHVTLKEYRFFN 180
Db          121 FQAVWSRALCSPPVCEAFARGNISSEEVCKPLCATEPRPGLAQACSSYSHVTLKEYRFFN 180
Qy          181 LOVLYPLISDPAINLRVHLYDRPRAVIRSRECTAKALARDNCIVLGNTGTMVEADERLR 240
Db          181 LOVLYPLISDPAINLRVHLYDRPRAVIRSRECTAKALARDNCIVLGNTGTMVEADERLR 240
Qy          241 VINEVCRSHVRIAREALKKPPPEQLDRYVRLVREYEDLARDPLTVIREYAFGIGLIPQLQ 300
Db          241 VINEVCRSHVRIAREALKKPPPEQLDRYVRLVREYEDLARDPLTVIREYAFGIGLIPQLQ 300
Qy          301 TWIHNITHGSGPGRAREFKTTSRDLSVQAMRHTLPFAKIRRVQELCGGALQIGYRS 360
Db          301 TWIHNITHGSGPGRAREFKTTSRDLSVQAMRHTLPFAKIRRVQELCGGALQIGYRS 360
Qy          361 VHSELEQRDLSDLPLPGMDSPKWASSTEKOPES 395
Db          361 VHSELEQRDLSDLPLPGMDSPKWASSTEKOPES 395
```

RESULT 2
 AAU11275
 ID AAU11275 standard; protein; 395 AA.
 XX
 AC
 XX
 DT 12-MAR-2002 (first entry)

DE Murine intestinal GlcNAc 6-sulfotransferase (I-GlcNAc6ST) protein.
 Qy 181 LQVLYPLISDPAINLRVHLYDRPRAVIRSREQTAKALARDNCIVLGNTGTMVEADPLRL 240
 Db 181 LQVLYPLISDPAINLRVHLYDRPRAVIRSREQTAKALARDNCIVLGNTGTMVEADPLRL 240
 Qy 1 MRLPRESSTVMSLIMVGTQSLIVELVSQVPSSPAGLGERVHVLVLSWRSGSSFVGOLF 60
 Db 1 MRLPRESSTVMSLIMVGTQSLIVELVSQVPSSPAGLGERVHVLVLSWRSGSSFVGOLF 60
 Qy 61 SQHDPVYFLMEPAWHDVTLISQSSAPALHMAYDLDLRSVFLCDMDVFDAYLPWRNISDL 120
 Db 61 SQHDPVYFLMEPAWHDVTLISQSSAPALHMAYDLDLRSVFLCDMDVFDAYLPWRNISDL 120
 Qy 121 FQAVWSRALCSPPVCEAFARGNISSEEVCKPLCATEPRPGLAQACSSYSHVTLKEYRFFN 180
 Db 121 FQAVWSRALCSPPVCEAFARGNISSEEVCKPLCATEPRPGLAQACSSYSHVTLKEYRFFN 180
 Qy 241 VVNEVCRSHVRIAREALKKPPPEQLDRYVRLVREYEDLARDPLTVIREYAFGIGLIPQLQ 300
 Db 241 VVNEVCRSHVRIAREALKKPPPEQLDRYVRLVREYEDLARDPLTVIREYAFGIGLIPQLQ 300
 Qy 301 TWIHNITHGSGPGRAREFKTTSRDLSVQAMRHTLPFAKIRRVQELCGGALQIGYRS 360
 Db 301 TWIHNITHGSGPGRAREFKTTSRDLSVQAMRHTLPFAKIRRVQELCGGALQIGYRS 360
 Qy 361 VHSELEQRDLSDLPLPGMDSPKWASSTEKOPES 395
 Db 361 VHSELEQRDLSDLPLPGMDSPKWASSTEKOPES 395.

OS Mus musculus.
 PN WO2001051177-A1.
 XX
 PD 15-NOV-2001.
 XX

PF 10-MAY-2001; 2001WO-US015452.
 XX
 PR 11-MAY-2000; 2000US-0056920.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Fukuda M, Yeh J, Hiraoka N;
 XX
 DR WPI; 2002-075226/10.
 XX
 PR New enzyme, useful for modifying acceptor molecule, comprises an isolated
 PR L-selectin sulfotransferase-2 that directs expression of L-selectin
 PR ligand antigen. MECA-79 in Chinese hamster ovary cells, or intestinal
 PR GlcNAc 6-sulfotransferase
 XX

Claim 28; Fig 10; 9pp; English.

PS The present invention provides a method of modifying an acceptor molecule
 XX by contacting the acceptor with an isolated betal-3-N-
 CC acetylglucosaminyltransferase (beral-3Gnt) or an active fragment, where
 CC betal-3Gnt directs expression of a MECA-79 antigen. The invention also
 CC provides a method of treating or preventing an L-selectin-mediated
 CC condition by reducing the expression or activity of a betal-3Gnt that
 CC directs expression of a MECA-79 antigen. This can be done by
 CC administering to the subject an oligosaccharide L-selectin antagonist
 CC that inhibits binding of L-selectin to a MECA-79 antigen, for example by
 CC administering antibody material that specifically binds betal-3Gnt,
 CC and/or a betal-3Gnt antisense nucleic acid molecule. L-selectin
 CC sulfotransferase-2 (LST-2) also directs MECA-79 antigen expression.
 CC Alternatively, the expression or activity of LST-2 or its active
 CC fragment can be reduced in combination with reducing the expression or
 CC activity of betal-3Gnt. The method is useful for treating L-selectin
 CC mediated conditions such as Crohn's disease and ulcerative colitis,
 CC inflammatory disorders of the skin such as allergic contact dermatitis,
 CC psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type
 CC hypersensitivity reactions, diabetes and hyperplastic thymus. This
 sequence represents mouse I-GlcNAc6ST
 XX Sequence 395 AA.

Query Match Score 2057; DB 5; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.7e-205;
 Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRLPRESSTVMSLIMVGTQSLIVELVSQVPSSPAGLGERVHVLVLSWRSGSSFVGOLF 60
 Db 1 MRLPRESSTVMSLIMVGTQSLIVELVSQVPSSPAGLGERVHVLVLSWRSGSSFVGOLF 60
 Qy 61 SQHDPVYFLMEPAWHDVTLISQSSAPALHMAYDLDLRSVFLCDMDVFDAYLPWRNISDL 120
 Db 61 SQHDPVYFLMEPAWHDVTLISQSSAPALHMAYDLDLRSVFLCDMDVFDAYLPWRNISDL 120
 Qy 121 FQAVWSRALCSPPVCEAFARGNISSEEVCKPLCATEPRPGLAQACSSYSHVTLKEYRFFN 180
 Db 121 FQAVWSRALCSPPVCEAFARGNISSEEVCKPLCATEPRPGLAQACSSYSHVTLKEYRFFN 180
 Qy 181 LOVLYPLISDPAINLRVHLYDRPRAVIRSRECTAKALARDNCIVLGNTGTMVEADERLR 240
 Db 181 LOVLYPLISDPAINLRVHLYDRPRAVIRSRECTAKALARDNCIVLGNTGTMVEADERLR 240
 Qy 241 VINEVCRSHVRIAREALKKPPPEQLDRYVRLVREYEDLARDPLTVIREYAFGIGLIPQLQ 300
 Db 241 VINEVCRSHVRIAREALKKPPPEQLDRYVRLVREYEDLARDPLTVIREYAFGIGLIPQLQ 300
 Qy 301 TWIHNITHGSGPGRAREFKTTSRDLSVQAMRHTLPFAKIRRVQELCGGALQIGYRS 360
 Db 301 TWIHNITHGSGPGRAREFKTTSRDLSVQAMRHTLPFAKIRRVQELCGGALQIGYRS 360
 Qy 361 VHSELEQRDLSDLPLPGMDSPKWASSTEKOPES 395
 Db 361 VHSELEQRDLSDLPLPGMDSPKWASSTEKOPES 395

| | |
|---|--|
| RESULT 3 ID ABB81557 XX | Qy 241 VNEYCVRSHVRIAALAHKPPPFQDPRLYRVEDLARDPLTVIRELYAFTGLGLTPOLQ 300 Db 265 VNBVCVSHVIAE-AHKPPPFQDPRLYRVEDLARDPLTVIRELYAFTGLGLTPOLQ 323 |
| AC | Qy 301 TWIHNTHGSGPARRAFKITSRDAISVSQAWRHTLPFAKIRRVQELCGGALQLQYRS 360 Db 324 TWIHNTHGSGPARRAFKITSRDAISVSQAWRHTLPFAKIRRVQELCGGALQLQYRS 383 |
| DT 05-SEP-2002 (first entry) | Qy 361 VISELEBODLSLULLRGMDSFKWASSTEKEOPES 395 Db 384 VHSELEORDLSLULLRGMDSFKWASSTEKEOPES 418 |
| XX | Mouse intestinal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:5. Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAcST; corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD; KW ophthalmological. |
| XX | RESULT 4 ID ABB81555 XX Standard; protein; 395 AA. AC ABB81555; XX DT 05-SEP-2002 (first entry) DB Consensus N-acetylglucosamine-6-sulfotransferase SEQ ID NO:3. XX Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAcST; KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD; KW ophthalmological. XX OS Homo sapiens. XX OS Mus musculus. OS Synthetic. |
| OS Mus musculus. PN US2002051562-A1. XX | Key PH Misc-difference 10 FT /label= Ala, Thr, Val FT Misc-difference 13 PT /label= Ala, Val, Ser FT Misc-difference 20 FT /label= Phe, Cys, Gly FT Misc-difference 39 FT /label= Ala, Asp, Glu FT Misc-difference 96 PT /label= Val, Met, Ile FT Misc-difference 142 FT /label= Ala, Thr, Asn FT Misc-difference 147 FT /label= Ala, Asp, Glu FT Misc-difference 159 FT /label= Thr, Ser, Gly FT Misc-difference 238 FT /label= Gly, His, Arg FT Misc-difference 294 FT /label= Ser, Thr, Gly FT Misc-difference 371 PT /label= Ala, Thr, Ser FT Misc-difference 380 PT /label= Leu, Pro, Met FT Misc-difference 382 FT /label= Gly, His, Ser FT Misc-difference 384 PT /label= Thr, Ser, Lys FT Misc-difference 390 FT /label= Ala, Glu FT Misc-difference 391 FT /label= Ser, Lys FT Misc-difference 392 PT /label= His, Gln FT Misc-difference 394 PT /label= Arg, Glu FT Misc-difference 395 PT /label= Asn, Ser XX PN US2002051562-A1. |
| XX | 23-MAY-2002. XX PF 09-AUG-2001; 2001US-00927602. XX PR 11-AUG-2000; 2000US-00638211. PR 11-AUG-2000; 2000US-0325773P. XX PA (FUKU//) FUKUDA M. N. PA (AKAM//) AKAMA T. O. PI Fukuda MN, Akama TO; XX DR 2002-507643/54. XX PT New nucleic acid encoding corneal N-acetylglucosamine-6-sulfotransferase, PT useful for treatment, monitoring and diagnosis of macular corneal PT dystrophy. XX PS Example 5; Page 24-25; 69pp; English. XX CC The present invention describes human corneal N-acetylglucosamine-6- CC sulfotransferase (I), which is able to catalyse sulfation of keratan CC sulfate (KS). Also described is a method for monitoring the effect of CC treatments for macular corneal dystrophy (MCD), and detecting CC susceptibility to MCD. (I) is located to chromosome 16q22, and has CC ophthalmological activity. (I) can be used to treat or prevent macular CC corneal dystrophy types I or II. (I) makes possible treatment of MCD CC without requiring keratoplasty or keratotomy. The present sequence CC represents mouse intestinal N-acetylglucosamine-6-sulfotransferase, CC which is given in comparison with (I) in the exemplification of the CC present invention XX SQ Sequence 418 AA; |
| Query Match Similarity 99.3%; Score 2042.5; DB 5; Length 418; Best Local Similarity 99.7%; Pred. No. 6e-204; Gaps 1; Matches 394; Conservative 0; Mismatches 0; Indels 1; | Query 1 MRLPRFSSTMVLSLIMVQTILFLVSRQPVSSPAAGLGRVHVLSSRSGSSFVGOLF 60 Db 25 MRLPRFSSTMVLSLIMVQTILFLVSRQPVSSPAAGLGRVHVLSSRSGSSFVGOLF 84 Query 61 SQHPDVYLMPEAWVWDTSIQSAPALMAVARDLIRSFLCDMDVFDAYLPWERNISDL 120 Db 85 SQHPDVYLMPEAWVWDTSIQSAPALMAVARDLIRSFLCDMDVFDAYLPWERNISDL 144 Query 121 FQWAVSRALCSPPVCEAFARGNISSEBVCPKLCATRPFGIAQEACSSYSHVILKEVRFFN 180 Db 145 FQWAVSRALCSPPVCEAFARGNISSEBVCPKLCATRPFGIAQEACSSYSHVILKEVRFFN 204 Query 181 IQVLYPLLSSPALNLRIVHYRDPAVLSREQTAKALARDNGIVLGNTGNTWADPRLR 240 Db 205 IQVLYPLLSSPALNLRIVHYRDPAVLSREQTAKALARDNGIVLGNTGNTWADPRLR 264 |

| | | | |
|---------|---|--|---|
| PI | Policky JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM; | Oy | 361 VHSELLEORDLSLDLILPRGMDPSKMASSTEKEOPES 395 |
| PI | Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS; | Db | 361 VYSEDEGRNLALDVLIFRGLNGFTWASSTASHPRN 395 |
| PI | Gandhi AR, Reddy R, Khan FA, Baugh MR, Ramkumar J, Griffin JA; | | |
| PI | Au-Young J; | | |
| XX | | | |
| DR | WPI: 2002-066363/09. | RESULT 8 | |
| DR | N-PSDB; AAD21670. | ADI21066 | |
| XX | Novel isolated human drug metabolizing enzymes referred as DME 1-10, useful for diagnosing, treating, or preventing disorders associated with aberrant expression of DME such as allergy, anemia, asthma, infertility. | ID | ADI21086 standard; protein; 395 AA. |
| PT | | AC | ADI21086; |
| PT | | XX | |
| PT | | DT | 15-APR-2004 (first entry) |
| XX | | XX | |
| PS | Claim 1a; Page 131-132; 143pp; English. | DE | Novel human protein #61. |
| XX | The invention relates to human drug metabolising enzymes referred as DME and nucleic acid molecules encoding such enzymes. Polynucleotides of the invention are useful for assessing toxicity of test compounds and in gene therapy. Sequences of the invention are useful in the diagnosis, prevention and treatment of autoimmune/inflammatory disorders such as acquired immune deficiency syndrome (AIDS), adult respiratory distress syndrome, allergies, anaemia, atherosclerosis, asthma, autoimmune haemolytic anaemia, contact dermatitis, Crohn's disease, giromerulonephritis, Goodpasture's syndrome, gout, Grave's disease, Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis, osteoarthritis, ostoporosis, psoriasis, systemic lupus erythematosus, rheumatoid arthritis, scleroderma, ulcerative colitis, uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic infections: cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis, duchenne and Becker dystrophy, hepatitis, cushing's syndrome, cancers, myelodysplastic syndrome, epilepsy; endocrine disorders such as disorders of the hypothalamus and pituitary, resulting from lesions such as primary brain tumours, adenomas, infarction associated with pregnancy, aneuysms, vascular malformations: eye disorders such as conjunctivitis, iritis, retinitis, glaucoma, pigmentosa ; metabolic disorders such as Addison's disease, cystic fibrosis, diabetes, goitre, glycogen storage diseases, hypercholesterolaemia, hyperthyroidism, hypoglycaemia, lipid myopathies, Menkes' syndrome, mannosidosis, obesity; gastrointestinal disorders such as dysphagia, gastric carcinoma, anaesthesia, nausea, gastroenteritis. | OS | Homo sapiens. |
| CC | hyperbilirubinaemia, emesis, cirrhosis, diarrhoea, jaundice, Reye's syndrome, peliosis hepatis, hepatic vein thrombosis and developmental disorders. The present sequence is human DME-5 protein | XX | |
| CC | | PN | W02003025148-A2. |
| CC | | XX | |
| CC | | PD | 27-MAR-2003. |
| CC | | XX | |
| CC | | PF | 19-SEP-2002; 2002WO-US029964. |
| CC | | XX | |
| CC | | PR | 19-SEP-2001; 2001US-0323719P. |
| CC | | PR | 13-SEP-2002; 2002US-003233139. |
| CC | | XX | |
| CC | | PA | (HYSE-) HYSEQ INC. |
| CC | | XX | |
| CC | | PI | Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao Qa, Wang J; |
| CC | | PI | Ghosh M, Xue AJ, Wehrman T, Zhou P, Drmanac RT, Wang D; |
| CC | | PI | Haley-Vicente D; |
| CC | | XX | |
| CC | | DR | WPI: 2003-354603/33. |
| CC | | DR | N-PSDB; ADI21802. |
| CC | | XX | New polynucleotides and secreted proteins, useful for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve tissue growth or regeneration, in wound healing, and in tissue repair and replacement. |
| CC | | PR | |
| XX | Sequence 395 AA; | XX | |
| Qy | Query Match 83.0%; Score 1708; DB 5; Length 395; | Claim 20; SEQ ID NO 337; 156pp; English. | |
| Qy | Best Local Similarity 83.0%; Pred. No. 4.8e-16; | XX | |
| Matches | Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0; | CC | The invention relates to an isolated polynucleotide encoding a polypeptide with biological activity. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders and other traits, to assess biodiversity, as nutritional sources or supplements. |
| Db | 1 MRLPFRSSVMTLSLVMVQTLVFLVSQVPSSPAGLGERHVLVLSWRSSSSFVQLF 60 | CC | The polynucleotides may also be used as molecular weight markers, chromosome markers or map related gene positions, or as an antigen to raise anti-DNA antibodies or elicit immune responses. The polypeptides are useful for raising antibodies, as markers for tissues in which the corresponding polypeptide is expressed, for re-engineering damaged or diseased tissues, for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and/or nerve tissue growth or regeneration, in wound healing, in tissue repair and replacement, in healing of burns, incisions and ulcers, and in treating cancer. The present sequence represents the amino acid sequence of a novel human protein. |
| Qy | 1 MWLPRVSSVATAVLLAQTFILLFLVSRPGSSPAGEARHVVLVLSWRSSSSFVQLF 60 | CC | SQ Sequence 395 AA; |
| Db | 1 SOHDVFVLMPEPAWHDITLSSQGAPAHMAYRDLIRSVFLCDMDVFDAYLPWRNISDL 120 | CC | |
| Qy | 61 SOHDVFVLMPEPAWHDITLSSQGAPAHMAYRDLIRSVFLCDMDVFDAYLPWRNISDL 120 | CC | |
| Db | 61 NQHDVFVLMPEPAWHDITLSSQGAPAHMAYRDLIRSVFLCDMDVFDAYLPWRNISDL 120 | CC | |
| Qy | 121 FOWAVSRAISCPYCEAFARGNSSEEVCKPLCAATRPGGLAQACSSYSHVVLKEYRFEN 180 | CC | |
| Db | 121 FOWAVSRAISCPYCEAFARGNSSEEVCKPLCAATRPGGLAQACSSYSHVVLKEYRFEN 180 | CC | |
| Qy | 181 LQVLYPLISDPAINLRVHLVRDPAVRSREQTAKALARNDIVIGTINGTWADPLR 240 | CC | |
| Db | 181 LQVLYPLISDPAINLRVHLVRDPAVRSREQTAKALARNDIVIGTINGTWADPLR 240 | CC | |
| Qy | 241 VNVBCRSRVIAEAALKPPFLQDRYRLVRYEDLARDPLTVIRELYAFTGLTGLTPQ 300 | CC | |
| Db | 241 VNVBCRSRVIAEAATKPPFLQDRYRLVRYEDLARDPLTVIRELYAFTGLTGLTPQ 300 | CC | |
| Qy | 301 TWIHNITHSGGPGRARREFKTTSERDALSYSVQMRHTLIPFAKTRVQELCGGALQLIGYRS 360 | CC | Query Match 83.0%; Score 1708; DB 7; Length 395; |
| Db | 301 AWIHNITHSGGPGRARREFKTTSERDALSYSVQMRHTLIPFAKTRVQELCGGALQLIGYRS 360 | CC | Best Local Similarity 83.0%; Pred. No. 4.8e-16; |

be useful in the field of pharmacogenomics, in particular for determining drug sensitivity and in treating breast cancer, hypervascular diseases, angiogenesis and scars in wound healing. The current sequence is that of a human protein tyrosine kinase biomarker protein of the invention.

XX Sequence 395 AA:

| | Query | Match | Score | Length | DB | No. | Pred. | CC | CC | CC | CC |
|-----------|--|---|--------|-------------|-------|-------------|-------------|-----------------------|--------|------------|-----------|
| | | | | | | | | Best Local Similarity | 83.0t; | 83.0t; | 4.8e-169; |
| | | | | | | | | Matches | 24; | Mismatches | 43; |
| | | | | | | | | Indels | 0; | Gaps | 0; |
| Db | 1 MMLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 61 SQHPDVYFVLMPEPAWHWDTLSQLGSAPAHMARDLIRSVPFLCDMDVFDAYLPWRNNSDL 120 | 1 MWLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | Best Local Similarity | 83.0t; | Pred. No. | 4.8e-169; |
| Dy | 61 NQHPDVYFVLMPEPAWHWDTLSQLGSATLNAVRDVLRSVPFLCDMDVFDAYLPWRNNSDL 120 | 1 MWLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | Matches | 24; | Mismatches | 43; |
| Dy | 121 FQWAVSRLCSPPPCEAFARGNISSEEVCKPLCATRPFGLQAECASSYSTVLRKEVRFPN 180 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | Indels | 0; | Gaps | 0; |
| Dy | 121 FQWAVSRLCSPPPCEAFARGNISSEEVCKPLCATRPFGLQAECASSYSTVLRKEVRFPN 180 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 181 LQVLYPLLSDPALNRIVHLVDRPVALRSREQTAKALARNDNGIVLTGNGWHEADPLRL 240 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 181 LQVLYPLLSDPALNRIVHLVDRPVALRSREQTAKALARNDNGIVLTGNGWHEADPLRL 240 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 241 VVNECRSHVRIAEAAHLHKPPPFQDRYFLVRYEDLARDPLTVIRELYAFTGLGTPOQLQ 300 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 241 VVNECRSHVRIAEAAHLHKPPPFQDRYFLVRYEDLARDPLTVIRELYAFTGLGTPOQLQ 300 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 301 TWINITHSGSGARREAFKTSRDSLVSQAWRHTLPPAKIRYQELCGCALQOLGYRS 360 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 301 TWINITHSGSGARREAFKTSRDSLVSQAWRHTLPPAKIRYQELCGCALQOLGYRS 360 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 361 VHSELLEQDLSLDLILPGRMDSFKWASSTEKKQPE 395 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | 361 VHSELLEQDLSLDLILPGRMDSFKWASSTEKKQPE 395 | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| Dy | RESULT 9 ADL61235 standard; protein: 395 AA. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| ADL61235; | 03-JUN-2004 (first entry) | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | Human tyrosine kinase biomarker carbohydrate sulphotransferase 6 protein. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | predictor set; protein tyrosine kinase; cytostatic; antiangiogenic; vasoconstrictive; pulmonary; pharmacogenomic; drug sensitivity; breast cancer; hypervascular disease; angiogenesis; wound healing scar; human; biomarker; carbohydrate sulphotransferase 6; enzyme. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | Homo sapiens. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | W02004020583-A2. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | 11-MAR-2004. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | 26-AUG-2003; 2003WO-US026491. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | 27-AUG-2002; 2002US-0406385P. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | (BRIM) BRISTOL-MYERS SQUIBB CO. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | Huang F., Han X., Reeves KA., Amler L., Fairchild CR., Lee FY.; Shaw P. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | WPI; 2004-239171/22. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | N-PSDB; ADL61098. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | New predictor sets with a plurality of polynucleotides and/or polypeptides whose expression pattern predicts cell response to a compound that modulates protein tyrosine kinase activity, useful in treating breast cancer. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | Claim 9; SEQ ID NO 159; 649pp; English. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | The invention relates to a novel predictor set comprising a plurality of polynucleotides and/or polypeptides whose expression pattern is predictive of the response of cells to treatment with a compound that modulates protein tyrosine kinase activity, or members of the protein tyrosine kinase pathway. The molecules of the invention demonstrate cytostatic, antiangiogenic, vasoconstrictive and vulnerability activities and may | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | 25-JAN-2001. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | 25-JUL-2000; 2000WO-US019741. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | 20-JUL-1999; 99US-0144694P. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | 13-JUN-2000; 2000US-0059828. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |
| CX | (REGG) UNIV CALIFORNIA. | 1 MRLPRFSSVTMVLSSLMMYOTGILYFLVSRVOPSSPAGLERVHVLVLSWSRGSSPFQOLF 60 | 83.0t; | Score 1708; | DB 8; | Score 1708; | Length 395; | | | | |

| | | | |
|----------|---|--|--|
| XX | WPI; 2003-845369/78. | PD | 05-SEP-2001. |
| DR | | XX | 07-JUL-2000; 2000EP-00114089. |
| XX | Identifying a mitochondrial target for drug screening assays and for treating diseases associated with altered mitochondrial function. | PF | 08-JUL-1999; 99JP-00194486. |
| PT | Specifically, it refers to a method for identifying proteins of the human heart mitochondrial proteome that are useful for drug screening assays, as well as therapeutic targets. The present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy of ragged red fibre syndrome (MERRF) or cancer. Accordingly, these compositions have neuroprotective, nootropic, antidiabetic, anticonvulsant, antiarthritic, osteopathic, ophthalmological and cytostatic activities. This polypeptide sequence is a human heart mitochondrial protein of the invention. | PR | 11-JAN-2000; 2000JP-0011874. |
| PT | | PR | 02-MAY-2000; 2000JP-00183765. |
| XX | | XX | |
| PS | Claim 1; SEQ ID NO 2211; 180pp; English. | PA | (HELI-) HELIX RES INST. |
| XX | This invention relates to novel mitochondrial targets that can be used for therapeutic intervention in treating a disease associated with altered mitochondrial function. Specifically, it refers to a method for identifying proteins of the human heart mitochondrial proteome that are useful for drug screening assays, as well as therapeutic targets. The present invention describes a method for identifying such proteins that can be used in the treatment of various diseases associated with altered mitochondrial function including diabetes mellitus, Huntington's disease, osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy of ragged red fibre syndrome (MERRF) or cancer. Accordingly, these compositions have neuroprotective, nootropic, antidiabetic, anticonvulsant, antiarthritic, osteopathic, ophthalmological and cytostatic activities. This polypeptide sequence is a human heart mitochondrial protein of the invention. | XX | XX |
| SEQ | Sequence 394 AA; | Query Match | Score 1333; DB 7; Length 394; |
| DB | 1 MWLRVSSATAAILAQFLFLFLVSRGPSSPAGGEARVHLVLSWRSSSFVSQLF 60 | Best Local Similarity | 78.1%; Pred. No. 7.2e-130; |
| QY | 1 MRLPRESSTVMLSHIMVTGILYFLVLSROQPSSPAGIGERVHVYLSSWRSGSSFVOLF 60 | Matches | Mismatches 257; Conservative 19; Indels 6; Gaps 1; |
| DB | 1 SOHPDVYLMPEAHWVWOTLSQSOSAPALHMAVDLIRESEVELCDMDVFDAYLPRRNTSDL 120 | Score | 1022.5; DB 4; Length 386; |
| DB | 61 NQHDVFYLMPEAHWVWOTLSQSQAATHLMAVDLIRESEVELCDMDVFDAYLPRRNSDL 120 | Best Local Similarity | 56.7%; Pred. No. 1.9e-97; |
| QY | 121 FQAVYRALCSPPVCEAFARGNISSEEVCKPLCATRPGGLAQACASSYSHVVLKEYRFN 180 | Matches | Mismatches 216; Conservative 42; Indels 11; Gaps 7; |
| DB | 121 FQAVYRALCSPPVCEAFRGNSSEEVCKPLCATRPGGLAQACASSYSHVVLKEYRFN 180 | Score | 1 NBLPRFESSTVMSLIMVQTGILVLP--LVSQVQPS-SPAQLGERVHVYLSSWRSGSSFV 56 |
| QY | 181 LQVLYPLUSDPAINLRVHLVRDRAVRSRECTAKALARNDNGIVLTGNTVVEADRLR 240 | Best Local Similarity | 56.7%; Pred. No. 1.9e-97; |
| DB | 181 LQVLYPLUSDPAINLRVHLVRDRAVRSRECTAKALARNDNGIVLTGNTVVEADRLR 240 | Matches | Mismatches 122; Indels 11; Gaps 7; |
| QY | 241 VNEVYCRSHVRIAEEAHLKPPFLQDRYRLVRYEDLARDPFLTVIRELYAFTGIGLTPLQO 300 | Score | 176 VRFENLQVLYPLUSDPAINLRVHLVRDRAVRSRECTAKALARNDNGIVLTGNTVVEADRLR 240 |
| DB | 241 LIREVCRSHVRIAEEAHLKPPFLQDRYRLVRYEDLARDPFLTVIRELYAFTGIGLTPLQO 300 | Best Local Similarity | 56.7%; Pred. No. 1.9e-97; |
| QY | 301 TWIHNITHCG-----PCGARREAFKTT 323 | Matches | Mismatches 119; Conservative 29; Indels 11; Gaps 7; |
| DB | 301 AWIHNITHGRGSASQSKPSILRLGMRTS 329 | Score | 179 VRFENLQVLYPLUSDPAINLRVHLVRDRAVRSRECTAKALARNDNGIVLTGNTVVEADRLR 240 |
| RESULT | RESULT 14 | QY | 116 NTSDLFQWASRALSCLSPPVCEAFARGNISSEEVCKPLCATRPGGLAQACASSYSHVVLKE 175 |
| AAM93309 | ID AAM93309 standard; protein; 386 AA. | DB | 119 RQSLFROWNSRALSCLSPACAD1IPQDE1IPRAHCRLUCSQQPPEVEXCRSYSHVVLKE 178 |
| XX | | QY | 176 VRFENLQVLYPLUSDPAINLRVHLVRDRAVRSRECTAKALARNDNGIVLTGNTVVEADRLR 240 |
| AC | | DB | 179 VRFENLQVLYPLUSDPAINLRVHLVRDRAVRSRECTAKALARNDNGIVLTGNTVVEADRLR 240 |
| XX | | QY | 235 ADPRLRVNEVERSRHVRIAEAALHKPPFLQDRYRLVRYEDLARDPFLTVIRELYAFTGIGLT 294 |
| DT | 06-NOV-2001 (first entry) | DB | 239 EDOPYTMQVICOSQLEYK-TIQSLKALQERYLLTRYEDLARAPTAQTSRMYEFVGLG 297 |
| XX | | QY | 295 LTPQLOTHWNIKTHGSPGARREAFKTSRDALSVQAWRHTLPFAKIRRVVOBLGGALQ 354 |
| DE | Human polypeptide, SEQ ID NO: 2817. | DB | 298 FLPHLQTWHNTRKGNG--DHAFFNWARDALNQWRSPPYEKVSRLQKAGDAMN 355 |
| XX | | QY | 355 LIGYRSHSELEFORDSLIDL 375 |
| KW | Human; full length cDNA; cDNA synthesis; oligo-capping. | DB | 356 LIGYRHRTSEQEORNLIDLL 376 |
| XX | Homo sapiens. | OS | RESULT 15 |
| XX | EP1130094-A2. | PN | ADL30784 |
| XX | | ID ADL30784 standard; protein; 386 AA. | |

| | | | | | |
|----|--|-----|-----|---|-----|
| AC | ADL30784. | Db | 298 | FLPHLQTWVHNITRGKGMG--DHAFHTNARDALNVSSQAWRWSLPYEKVSRQLQACGGDMN | 355 |
| XX | DT 20-MAY-2004 (first entry) | Qy | 355 | LIGYRSTHSELEBORDLSDLL | 375 |
| XX | Human protein encoded by a full length cDNA clone SegID 2817. | Db | 356 | LIGYRHRSEQEORNLIDLL | 376 |
| XX | human; medicine; signal transduction; glycoprotein; transcription; | | | | |
| KW | oligo-capping method. | | | | |
| XX | Homo sapiens. | | | | |
| OS | XX | | | | |
| XX | EP1396543-A2. | | | | |
| PN | | | | | |
| XX | PD 10-MAR-2004. | | | | |
| XX | PP 07-JUL-2000; 2003BP-00025638. | | | | |
| PR | 08-JUL-1999; 99JJP-00194486. | | | | |
| PR | 11-JAN-2000; 20000JP-00118774. | | | | |
| PR | 02-MAY-2000; 20000JP-00183865. | | | | |
| PR | 07-JUL-2000; 20000BP-00114089. | | | | |
| XX | (REAS-) RES ASSOC BIOTECHNOLOGY. | | | | |
| XX | PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y; | | | | |
| PI | PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H; | | | | |
| XX | DR WPI, 2004-204755/20. | | | | |
| DR | N-PSDB; ADL30783. | | | | |
| XX | PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full | | | | |
| PT | length human cDNAs. | | | | |
| PT | PS XX Example 1; SEQ ID NO 2817; 1340pp; English. | | | | |
| XX | CC This invention relates to a novel primers useful for synthesising full | | | | |
| CC | CC length cDNA molecules that encode human proteins. Specifically, it refers | | | | |
| CC | CC to secretory or membrane proteins that are potential therapeutic agents/ | | | | |
| CC | CC target molecules in the field of medicine, and in particular genes | | | | |
| CC | CC encoding proteins that are associated with signal transduction, | | | | |
| CC | CC glycoproteins and transcription. The present invention describes a method | | | | |
| CC | CC for efficiently cloning a full length human cDNA from both the 5' and 3' | | | | |
| CC | CC ends using the oligo-capping method. This polypeptide sequence is a full | | | | |
| XX | CC length human protein of the invention. | | | | |
| XX | SQ Sequence 386 AA; | | | | |
| Qy | Query Match 49.7%; Score 1022.5; DB 8; Length 386; | | | | |
| Qy | Best Local Similarity 56.7%; Pred. No. 1.9e-97; | | | | |
| Qy | Matches 216; Conservative 42; Mismatches 112; Indels 11; Gaps 7; | | | | |
| Db | 1 MRLPRFSSTMISLMLVQTCILVP--LIVSRQVPS-SPAGLGERVHVLVLSVRSRGSSFV | 56 | | | |
| Db | 1 MLLPK- KKK MLFLVSVQNLALFFHMYSHNTISLSMAQPERNMHLVLSVRSRGSSFV | 58 | | | |
| Qy | 57 GOLFSQHPDVFYLMEPAWHTWDTSIQSAPALMNAVRDLIRSVLCDMDVFDAYL-PVRR | 115 | | | |
| Db | 59 GQLFQHPDVFYLMEPAWHTWMTPQSTAWMLNAVRDLIRAYFLCDMSVFDAYMPEPQR | 118 | | | |
| Qy | 116 NISDLFQWAVSRAICGPPVCEAFARGNISSVECKPLCATRPFGLAQEACSSYSHVLIK | 175 | | | |
| Db | 119 RQSSLQWENSRAICGPPVCEAFARGNISSVECKPLCATRPFGLAQEACSSYSHVLIK | 178 | | | |
| Qy | 176 VRFENLQVLYPLSPDLPALNRIVHLYRDPRAVLRSRECTAKALARNDGTVLGNITW-E | 234 | | | |
| Db | 179 VRFENLQSLYPLLKDPNSLNHLYRDPRAVLRSRECTAKALARNDGTVLGNITW-E | 238 | | | |
| Qy | 235 ADPRLRVNNVCRSHVRIAEAALHKPPFLQDRYRLVRYEDLARDPLTVIRELYAFTGLG | 294 | | | |
| Db | 239 EDQPYMMQVICQSQLEIYK-TIQSLPKAQLQERVILVRYEDLARPAQTSRMYEFVGL | 297 | | | |
| Qy | 295 LTPQLOQTWINTHGSGPGRARAEFKTTSRDALSYSQAWRHTLPFAKIRROELCGGALQ | 354 | | | |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model
Run on: June 23, 2005, 08:33:38 ; Search time 8.95931 Seconds
(without alignments)
4242.027 Million cell updates/sec

Title: US-10-697-828-7
Perfect score: 2057
Sequence: .1 MRLPRSSSTVNLSLMVQTG.....LPRGMDSFKWASSTEKQPES 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters:

283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_79;*
1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 643.5 | 31.3 | 484 | 2 JE0261 | N-acetylglucosamin |
| 2 | 580 | 28.2 | 486 | 2 JC7351 | N-acetylglucosamin |
| 3 | 577 | 28.1 | 484 | 2 JC7350 | N-acetylglucosamin |
| 4 | 495 | 24.1 | 458 | 2 A57397 | chondroitin 6-sulf |
| 5 | 148.5 | 7.2 | 307 | 2 B95934 | probable enzyme, C |
| 6 | 126 | 6.1 | 388 | 2 G70729 | hypothetical prote |
| 7 | 101.5 | 4.9 | 262 | 2 TS5999 | Probable aminoglyc |
| 8 | 99 | 4.8 | 571 | 2 AG3174 | acetolactate synth |
| 9 | 98 | 4.8 | 856 | 2 B87316 | conserved hypothet |
| 10 | 95 | 4.6 | 547 | 2 A64878 | peptide transport |
| 11 | 94.5 | 4.6 | 892 | 1 S42228 | replication licens |
| 12 | 93 | 4.5 | 547 | 2 B85756 | hypothetical prote |
| 13 | 93 | 4.5 | 547 | 2 B85756 | hypothetical prote |
| 14 | 93 | 4.5 | 559 | 2 T03412 | malate synthase (E |
| 15 | 92.5 | 4.5 | 327 | 2 T0744 | spheroiodene monoox |
| 16 | 92.5 | 4.5 | 4077 | 2 T17484 | hypothetical prote |
| 17 | 91 | 4.4 | 562 | 2 T48413 | malate synthase-1 |
| 18 | 91 | 4.4 | 925 | 2 T02811 | DNA excision/repair |
| 19 | 90 | 4.4 | 933 | 1 DEECOG | oxoglutarate dehyd |
| 20 | 90 | 4.4 | 933 | 2 F85573 | hypothetical prote |
| 21 | 90 | 4.4 | 933 | 2 G90722 | oxoglutarate dehyd |
| 22 | 89.5 | 4.4 | 311 | 2 B40216 | flavonol 3'-sulfot |
| 23 | 89.5 | 4.4 | 459 | 2 G83174 | exodooxyribonuclea |
| 24 | 88 | 4.3 | 668 | 2 T01685 | cprotein - mai |
| 25 | 87.5 | 4.3 | 342 | 2 T19687 | lipoyltransferase |
| 26 | 87.5 | 4.3 | 678 | 2 T23341 | beta-catenin - Cae |
| 27 | 87.5 | 4.3 | 860 | 2 F90854 | mannosyltransferas |
| 28 | 87.5 | 4.3 | 1158 | 2 F90854 | probable host spec |
| 29 | 87 | 4.2 | 413 | 2 H86825 | hypothetical prote |

ALIGNMENTS

| | | |
|---|---|---|
| RESULT 1 | JE0261 | N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human |
| A;Species: | Homo sapiens (man) | C;Comment: This protein catalyzes the transfer of sulfate from 3'-phosphoadenosine 5'-phosphate to N-acetylglycosamine. |
| C;Accession: | JE0261 | C;Superfamily: chondroitin 6-sulfotransferase |
| R.Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, Q.W.; Mitsuo, T. | J. Biochem. 124, 670-678, 1998 | A;Reference number: JE0261; PMID:98391845; PMID:9722682 |
| A;Accession: JE0261 | A;Molecule type: mRNA | A;Cross-references: DDBJ:AB014679 |
| A;Residues: 1-484 <UCH> | Best Local Similarity 36.6%; Pred. No. 8.3e-48; Mismatches 77; Indels 8 Gaps 9; | C;Keywords: sulfotransferase |
| SPAGIGERH-VVILSNSRSGSSSTVGQLSPQHPDVYMEPAFTWDTLSQGSAPALHMA 91 | SPAGIGERH-VVILSNSRSGSSSTVGQLSPQHPDVYMEPAFTWDTLSQGSAPALHMA 91 | Query Match 31.3%; Score 643.5%; DB 2; |
| 110 APEGVGDKHMMVFTWNSGSSFEGELQNPNPEVFLBPWMWVQKLYPGDAVSLOGA 169 | 110 APEGVGDKHMMVFTWNSGSSFEGELQNPNPEVFLBPWMWVQKLYPGDAVSLOGA 169 | Best Local Similarity 36.6%; Pred. No. 8.3e-48; Mismatches 77; Indels 8 Gaps 9; |
| 92 VRDLIRSVPLCDMDVFDAYLP--WRWNISDL--FQWAVSRALCSPCPVCEAFARG-NIS 144 | 92 VRDLIRSVPLCDMDVFDAYLP--WRWNISDL--FQWAVSRALCSPCPVCEAFARG-NIS 144 | Matches 139; Conservative 77; Mismatches 117; Indels 8 Gaps 9; |
| 170 ARDMSALYRCDSLSPAGSGGRNLTTLG1FGAATNKWVCSSPLCPAYKEVVGVLV 229 | 170 ARDMSALYRCDSLSPAGSGGRNLTTLG1FGAATNKWVCSSPLCPAYKEVVGVLV 229 | Query Match 31.3%; Score 643.5%; DB 2; |
| 289 RAVASSRIRSRRGLIRESQVVRER-----DPAHMRPPEAAGHKLGAKKKGIVGPAD 342 | 289 RAVASSRIRSRRGLIRESQVVRER-----DPAHMRPPEAAGHKLGAKKKGIVGPAD 342 | Best Local Similarity 36.6%; Pred. No. 8.3e-48; Mismatches 77; Indels 8 Gaps 9; |
| 343 YHALGAMETVICNSMAKLTQTL-QPPDWLQGHYLVRYEDLVLGEFPKTURRYDPGLLV 401 | 343 YHALGAMETVICNSMAKLTQTL-QPPDWLQGHYLVRYEDLVLGEFPKTURRYDPGLLV 401 | Matches 139; Conservative 77; Mismatches 117; Indels 8 Gaps 9; |
| 296 TPOLOTWIINITHSSGGARAREAKFTSDALSYSQAWRHTLPPAKIRRQBLCCGALQL 355 | 296 TPOLOTWIINITHSSGGARAREAKFTSDALSYSQAWRHTLPPAKIRRQBLCCGALQL 355 | Query Match 31.3%; Score 643.5%; DB 2; |
| 402 SPENMQFALNMTSQSSSSSK - PVVUSARNQANWTAFTQQIKQVEEFCYOPMVA 459 | 402 SPENMQFALNMTSQSSSSSK - PVVUSARNQANWTAFTQQIKQVEEFCYOPMVA 459 | Matches 139; Conservative 77; Mismatches 117; Indels 8 Gaps 9; |
| 356 LGYRSVSELEQORDLSLDL 375 | 356 LGYRSVSELEQORDLSLDL 375 | Query Match 31.3%; Score 643.5%; DB 2; |
| 460 LGYEVNSPEEVKDLKSTLL 479 | 460 LGYEVNSPEEVKDLKSTLL 479 | Matches 139; Conservative 77; Mismatches 117; Indels 8 Gaps 9; |

RESULT 2
 JC7351 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - human
 C: Species: Homo sapiens (man)
 C: Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
 C: Accession: JC7351 R: Uchimura, K.; Pasakhany, F.; Kadomatsu, K.; Yamakawa, T.; Kurosoawa, N.; Biochem. Biophys. Res. Commun. 274, 291-296, 2000
 A: Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
 A: Reference number: JC7350; MUID:20374462; PMID:10913333
 A: Accession: JC7351
 A: Molecule type: mRNA
 A: Residues: 1-486 <UCH>
 A: Cross-references: UNIPROT:O75667; DDBJ:AB040711
 C: Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in
 C: Keywords: glycolysis; sulfotransferase; transmembrane protein; sulfate transport; sulfotransferase; transmembrane protein

Query Match Score 580; DB 2; Length 486;
 Best Local Similarity 34.8%;保守性 56%; Indels 66; Gaps 8;
 Matches 139; Conservative 56%; Mismatches 137; Indels 66; Gaps 8;

Qy 4 PRFSTVNLIMVOTGILIVFLYRSQVEPSSPAGLGERVHVLVLSWRSGSSFVGOLFQH 63
 Db 83 PRFPNL-----SGAVGEAVSR----- 124

Qy 64 PDVFYLMPAWHVDLTSQGASALMNAVRDILRSVFLCDMDVFDAYLP----- 112
 Db 125 PDVFYLMPWMLWQALYPGDAESLGQLDRMSPRLRCDFSVLRLYAPPGDPAARAPDT 184

Qy 113 WRRNISDLFQWAWSRALCSPV----EAFARGNISSEBEVKPLCATRPFGLAQEACSSY 168
 Db 185 ANLTTAALFRWRPNKVICSPPLCPGAPARAEVGLVETACERSCPVVAIRALEBCKY 244

Qy 169 SHVLFKEYRFNQLVLYPLSDPALNRLRVHVLVDRPAVLRSREOTAKALARDNGIVL-- 226
 Db 245 PVVVKDVLQVFLRPSGLVPLRDPGLNURKVLQFRDPAVHNRSRLKSROGLLRESIOVLR 304

Qy 227 -----GTGNTGWVADPR-----LAVVNEVCRSHVRIAZAHL 258
 Db 305 RQRGDRFRHVRVLLAHGVGARGPQQSRLPAAPADFFELTGALVECEAWLDLIFARGA-- 362

Qy 259 KPPFLQDRYRLVRYEDIARDPLTVIRELYAFTGLGLTPQLQWVHINITHGGP GARREA 318
 Db 363 -PAWLRRYLRRLRYEDLVRQPAQLRLLRFSGLRALAALDAFALNMTRGAYGADR-P 419

Qy 319 EKFTSRDALSVSQAWRHTLPFAKIRYQELCGGALQIY 358
 Db 420 FHLSARDAREAVHAWRERLSRECVRQVEAACAPMRLLAY 459

RESULT 4
 Qy 113 WRRNISDLFQWAWSRALCSPV----EAFARGNISSEBEVKPLCATRPFGLAQEACSSY 168
 Db 185 ANLTTAALFRWRPNKVICSPPLCPGAPARAEVGLVETACERSCPVVAIRALEBCKY 244

Qy 169 SHVLFKEYRFNQLVLYPLSDPALNRLRVHVLVDRPAVLRSREOTAKALARDNGIVL-- 226
 Db 245 PVVVKDVLQVFLRPSGLVPLRDPGLNURKVLQFRDPAVHNRSRLKSROGLLRESIOVLR 304

Qy 227 -----GTGNTGWVADPR-----LAVVNEVCRSHVRIAZAHL 258
 Db 305 RQRGDRFRHVRVLLAHGVGARGPQQSRLPAAPADFFELTGALVECEAWLDLIFARGA-- 362

Qy 259 KPPFLQDRYRLVRYEDIARDPLTVIRELYAFTGLGLTPQLQWVHINITHGGP GARREA 318
 Db 363 -PAWLRRYLRRLRYEDLVRQPAQLRLLRFSGLRALAALDAFALNMTRGAYGADR-P 419

Qy 319 EKFTSRDALSVSQAWRHTLPFAKIRYQELCGGALQIY 358
 Db 420 FHLSARDAREAVHAWRERLSRECVRQVEAACAPMRLLAY 459

RESULT 3
 JC7350 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse
 C: Species: Mus musculus (house mouse)
 C: Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
 C: Accession: JC7350 R: Uchimura, K.; Pasakhany, F.; Kadomatsu, K.; Yamakawa, T.; Kurosoawa, N.; Biochem. Biophys. Res. Commun. 274, 291-296, 2000
 A: Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
 A: Reference number: JC7350; MUID:20374462; PMID:10913333
 A: Accession: JC7350
 A: Molecule type: mRNA
 A: Residues: 1-484 <UCH>
 A: Cross-references: UNIPROT:O9BP78; DDBJ:AB040710
 C: Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in
 C: Keywords: glycolysis; sulfotransferase; transmembrane protein; sulfate transport; sulfotransferase; transmembrane protein

Query Match Score 577; DB 2; Length 484;
 Best Local Similarity 34.5%;保守性 57%; Indels 58; Gaps 8;
 Matches 135; Conservative 57%; Mismatches 141; Indels 58; Gaps 8;

Oy 31 PSSPAGGG-----BRVHVVLSSWRSGSSFVGOLFQHOPDVFLMPEAWHVNDDTLSQ 82
 Db 82 PGSPGNISAVGEAVTQBKQHIYVHATWRTGSFLGELFNQHPDVFLYEPWHLHQALYP 141

Oy 83 GSAPALHMAYVDRDLIRSVFLCDMDVFDAY-----LPWRRNISD-LFWQAVSRALCS 131
 Db 142 GDAEISQALRDLMLRSFLRCDFSVLRYAQFGDPGERAPSANLTAMPLFRWTRNVC 201

Oy 132 PPVCEAFAARG---NISSEEVCKPLCATRPFGLAQACASSYSHVYLKEVRFNQLQVYPL 187
 Db 202 PPLCPAPARADAVGLVEDRAESTCPVSLRALLAECCRYPVNUKDVRLGVLVPL 261

Oy 188 LSDPAINLRLTHVLDRAVRSREOTAKALARDNGIVLGN-----GRWEAD 236
 Db 262 LRDPGLNLKVVLQFLRPAVNSRLKSROGLLRESIQVLRQGDHFHRVLLAIGVDA 321

Oy 237 P-----RLRVVNEVCRSHVRIAEEAALHKPKPPLQDRYRLVRYEDLA 277
 Db 322 PGQQAFLPASPRADEFFLTSALEVICEAWLRLDPLLTRGA---PAWLRRYLRRLVRYEDLV 377

Oy 278 RDPLTVIRELYAFTGLGLTPOLQTWHINITHGSGCARRAFKTTSRDALS VSQAWRHTL 337
 Db 378 WQPQAQIRRLRFSGLRTLALDAFAFMTRGSAVGDRL-PFHLSARDAREAVHWRERL 436

Oy 338 PFAKTRVQEIJGGALQQLGGYRSVHSEQR 368
 Db 437 SQEQVROVETACAPAMLLYPRSGDERRK 467

Oy 439 LRSREOTAKALARDNGIVLGN-----EAFAF 477

Oy 440 RVHVVLSSWRSGSSFVGOLFQHOPDVFLMPEAWHVNDDTLS--QGSAPALHMAY--RDL 95
 Db 112 RRHVLMLMTATRGSSVGEFFNOQNFYLFEPLWHLIERTVTFEPGGANAVGSALVYRDV 171

Oy 446 IRSVFLCDMDVFDAYL--PWRNISDLPVAVSRALCSPVVC----EAFAFGNISSEE 147
 Db 172 LQQLLUDLTLYESFSSPAPBEHLTAFLRGSSHSCLCBPVCTPSLKVFEKTHCKNRR 231

Oy 148 VCKPLCATRPFGLAQACASSYSHVYLKEVRFNQLQVYPLSDPAINLRLVHYDRPRAV 207
 Db 232 -CGPANT----LAAPACRKRQHMAKTVRQLEFLQPAIADPRDLRITQLVDRPRAV 286

Oy 208 LRSREOTAKALARDNGIVLGN-----VEADPRLRVNEV----CRSHVRIAEAL 257
 Db 287 LVSRWYAFS-----GYYESKKWAAEGEAPIQ-EDEVQRLRGNCSEIRLSAELGL 335

Oy 258 HKPPPLQDRYRLVRYEDLARDPLTVIRELYAFTGLGLTPOLQWTHNITHGSGPBARRE 317
 Db 336 -RQPWRMLGRMLVRYEDVARPLKALEMYRFA吉HPTPOVEWIRANTQAPQDS--N 391

Oy 318 AFKTSRDALSVSQAWRHTLPFAKIRYQELCGGALQIY 372

| |
|--|
| Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. |
| Nature 393, 537-544; 1998 |
| A; Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. |
| A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome |
| A; Reference number: A70500; PMID: 9634230 |
| A; Accession: G70729 |
| A; Status: preliminary; nucleic acid sequence not shown; translation not shown |
| A; Molecule type: DNA |
| A; Residues: 1-388 <COL> |
| A; Cross-references: UNIPROT:Q50695; GB:Z77163; ALI123456; NID:93261610; PIDN: CAB00968; A; Experimental source: strain H37Rv |
| C; Genetics: |
| A; Gene: Rv2267c |
| Query Match Score 126; DB 2; Length 388; |
| Best Local Similarity 23.1%; Pred. No. 0.0046; Gaps 17; |
| Matches 79; Conservative 50; Mismatches 131; Indels 82; |
| Qy 43 VLVSSMRSGSSPGQLP---SQRH---PPDVFYLMPEPAWHDITLSQGSAPALHMAVRLIR 97 |
| Db 83 IPIVGHRRTGTTLHELLVVDDBHTGPYCECLAPHHFL---LTEWFAPVE----- 131 |
| Qy 98 SVPLCD---MVDYFDAYLPWRRNISDLFOWAVSRALCSPPVCEBAFARGNISSEEVCKPLC 153 |
| Db 132 --FLVSKHRAMDMDF--LSSLHHHQDEFPWCM--OGLPSPLVTLAEPNRPPQQEY----- 181 |
| Qy 154 ATRPFGLAQEACSSYSHVLUKEVRPNLQLVY----PLSDPRLNRL----- 197 |
| Db 182 ----LDLQQVAPRLEFIWKRLPFRQVYPRRKTVLKNPHTSPRKVLLFVFPQA 235 |
| Qy 198 --VHLVRDPRAYERSRECTAKALARNDNGIVLGNTNGTWEDPRLRVNECRSHYRIAEA 255 |
| Db 236 KTHIIVRPYVYPSTHLKALYIRHGL---OOPTFDGLDD---KVVSFTVDSLRLKLDEG 290 |
| Qy 256 ALHKPPPFQDRYRVLRYVEDLADPLTVIRELYNAFTGLG---LTPQIQTWHINNTHGSG 311 |
| Db 291 RELVDP---TREFYELRYEDLICGPEGOLRRLYQHGLGDPECYPLRLROYLAD--HAD- 343 |
| Qy 312 PGARREAFPKTTSRDAISYSQAW-----RHTLPPFAKIR 343 |
| Db 344 --YKTNSQLTVEQARIAVDEHNGEIIDRYGYDHTPEPARL 383 |
| RESULT 7 |
| T35999 probable aminoglycoside acetyltransferase - Streptomyces coelicolor |
| C; Species: Streptomyces coelicolor |
| C; Date: 03-Dec-1999 #Sequence_revision 03-Dec-1999 #text_change 09-Jul-2004 |
| C; Accession: T35999 |
| R; Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. |
| submitted to the EMBL Data Library, July 1999 |
| A; Reference number: 221574 |
| A; Accession: T35999 |
| A; Status: Preliminary; translated from GB/EMBL/DBJ |
| A; Molecule type: DNA |
| A; Residues: 1-262 <SEQ> |
| A; Cross-references: UNIPROT:Q9XAC9; EMBL:AL096839; PIDN: CAB00752.1; GS PDB: GN00070; SCOPD |
| A; Experimental source: strain A3 (2) |
| C; Genetics: |
| A; Gene: SCOPDB:SCC22_09 |
| C; Superfamily: Escherichia coli aminoglycoside N3'-acetyltransferase |
| Query Match Score 101.5; DB 2; Length 262; |
| Best Local Similarity 29.4%; Pred. No. 0.37; Gaps 41; |
| Matches 47; Conservative 19; Mismatches 53; Indels 41; |
| Qy 225 VLGTGTTVNEADRRLRVNECRSHVRIAEEALHKPPPFQDRYRVLRYEDLARDPLRTVI 284 |
| Db 57 VLGPDTGLVY-----VPTQGDSPLWTSPVYBEEWWETVRAATPAPYDPL--- 102 |
| Qy 285 REIYATGIGLTPO--LQTWIINHTGSGIGARFKTSRDAISYSQAWRHTLPPAKIR 343 |

| | | | | |
|--|-----|---|-----|---|
| Db | 103 | -ITPSRGVGVVPETVRTW-----PGARRSAHPQTSAALGG-----R | 138 | A;Accession: F87316 A;Species: Agrobacterium tumefaciens C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 16-Aug-2004 C;Accession: AG3174 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Gerde, G.; Gillett, W.; Grant, C.; Guenther, D.; Kurayvin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S. A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, Stern, E.W. A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58. A;Reference number: AB5777; MUID:21608550; PMID:11743193 A;Accession: AG3174 A;Molecule type: DNA A;Cross-references: UNIPROT:Q8URH9; GB:AE008687; PIDN:AAU45813.1; PID:917743551; GSPPDB:G A;Experimental source: strain C58 (Dupont) C;Genetics: |
| Qy | 344 | RQECLCGNQL-LGYRSYHSELEFQDLSLDLJLRGMDS 362 | | |
| Db | 139 | AAEVVAGHATDCRGERSPAALERLDARV-LULGAGYDA 177 | | |
| RESULT 8 | | | | |
| AG3174 | | acetolactate synthase [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid | | |
| C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 16-Aug-2004 | | | | |
| C;Accession: AG3174 | | | | |
| R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Gerde, G.; Gillett, W.; Grant, C.; Guenther, D.; Kurayvin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S. | | | | |
| A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, Stern, E.W. | | | | |
| A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58. | | | | |
| A;Reference number: AB5777; MUID:21608550; PMID:11743193 | | | | |
| A;Accession: AG3174 | | | | |
| A;Molecule type: DNA | | | | |
| A;Cross-references: 1-571 <KUR> | | | | |
| A;Experimental source: strain C58 (Dupont) | | | | |
| C;Genetics: | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches | | | | |
| Conservative | | | | |
| Score | | | | |
| Length | | | | |
| Pred. No. | | | | |
| Mismatches | | | | |
| Indels | | | | |
| Gaps | | | | |
| 4.8% | | | | |
| DB 2; | | | | |
| Length 856; | | | | |
| A;Accession: F87316 | | | | |
| A;Species: preliminary | | | | |
| A;Molecule type: DNA | | | | |
| A;Residues: 1-856 <STO> | | | | |
| A;Cross-references: UNIPROT:Q9AAQ3; GB:AE005673; PIDN:AAK22530.1; PIDN:AK22530.1; GSPPDB:G | | | | |
| A;Gene: CC0543 | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches | | | | |
| Conservative | | | | |
| Score | | | | |
| Length | | | | |
| Pred. No. | | | | |
| Mismatches | | | | |
| Indels | | | | |
| Gaps | | | | |
| 19.5% | | | | |
| DB 2; | | | | |
| Length 856; | | | | |
| A;Accession: F87316 | | | | |
| A;Species: Caulobacter crescentus | | | | |
| C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 | | | | |
| C;Accession: F87316 | | | | |
| B;Laub, M.T.; DeBoy, R.R.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. | | | | |
| A;Title: Complete Genome Sequence of Caulobacter crescentus. | | | | |
| A;Reference number: A87249; MUID:21173698; PMID:11259647 | | | | |
| RESULT 9 | | | | |
| F87316 | | conserved hypothetical protein CC0543 [imported] - Caulobacter crescentus | | |
| C;Species: Caulobacter crescentus | | | | |
| C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 | | | | |
| C;Accession: F87316 | | | | |
| B;Laub, M.T.; DeBoy, R.R.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. | | | | |
| A;Title: Complete Genome Sequence of Caulobacter crescentus. | | | | |
| A;Reference number: A87249; MUID:21173698; PMID:11259647 | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches | | | | |
| Conservative | | | | |
| Score | | | | |
| Length | | | | |
| Pred. No. | | | | |
| Mismatches | | | | |
| Indels | | | | |
| Gaps | | | | |
| 22.0% | | | | |
| DB 2; | | | | |
| Length 547; | | | | |
| A;Accession: F87316 | | | | |
| A;Species: Caulobacter crescentus | | | | |
| C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 | | | | |
| C;Accession: F87316 | | | | |
| B;Laub, M.T.; DeBoy, R.R.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. | | | | |
| A;Title: Complete Genome Sequence of Caulobacter crescentus. | | | | |
| A;Reference number: A87249; MUID:21173698; PMID:11259647 | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches | | | | |
| Conservative | | | | |
| Score | | | | |
| Length | | | | |
| Pred. No. | | | | |
| Mismatches | | | | |
| Indels | | | | |
| Gaps | | | | |
| 27; | | | | |
| DB 2; | | | | |
| Length 466; | | | | |
| A;Accession: F87316 | | | | |
| A;Species: Caulobacter crescentus | | | | |
| C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 | | | | |
| C;Accession: F87316 | | | | |
| B;Laub, M.T.; DeBoy, R.R.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. | | | | |
| A;Title: Complete Genome Sequence of Caulobacter crescentus. | | | | |
| A;Reference number: A87249; MUID:21173698; PMID:11259647 | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches | | | | |
| Conservative | | | | |
| Score | | | | |
| Length | | | | |
| Pred. No. | | | | |
| Mismatches | | | | |
| Indels | | | | |
| Gaps | | | | |
| 162; | | | | |
| DB 2; | | | | |
| Length 27; | | | | |
| A;Accession: F87316 | | | | |
| A;Species: Caulobacter crescentus | | | | |
| C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 | | | | |
| C;Accession: F87316 | | | | |
| B;Laub, M.T.; DeBoy, R.R.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. | | | | |
| A;Title: Complete Genome Sequence of Caulobacter crescentus. | | | | |
| A;Reference number: A87249; MUID:21173698; PMID:11259647 | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches | | | | |
| Conservative | | | | |
| Score | | | | |
| Length | | | | |
| Pred. No. | | | | |
| Mismatches | | | | |
| Indels | | | | |
| Gaps | | | | |
| 170; | | | | |
| DB 2; | | | | |
| Length 27; | | | | |
| A;Accession: F87316 | | | | |
| A;Species: Caulobacter crescentus | | | | |
| C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 | | | | |
| C;Accession: F87316 | | | | |
| B;Laub, M.T.; DeBoy, R.R.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. | | | | |
| A;Title: Complete Genome Sequence of Caulobacter crescentus. | | | | |
| A;Reference number: A87249; MUID:21173698; PMID:11259647 | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches | | | | |
| Conservative | | | | |
| Score | | | | |
| Length | | | | |
| Pred. No. | | | | |
| Mismatches | | | | |
| Indels | | | | |
| Gaps | | | | |
| 61; | | | | |
| DB 2; | | | | |
| Length 46; | | | | |
| A;Accession: F87316 | | | | |
| A;Species: Caulobacter crescentus | | | | |
| C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004 | | | | |
| C;Accession: F87316 | | | | |
| B;Laub, M.T.; DeBoy, R.R.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.J.; Brmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. | | | | |
| A;Title: Complete Genome Sequence of Caulobacter crescentus. | | | | |
| A;Reference number: A87249; MUID:21173698; PMID:11259647 | | | | |

| | | | | | | | |
|----|-----|---|------------------------|--|---|--|-----|
| Qy | 47 | SSWRSGSSFYVGOLFQSOHQDWFYLMEPAMHWDTLSQGSAPALHMAVRDL----- | 95 | Db | 596 | QTSISIKAGIVTSLQARCTVIAANPPIGGRYDPSELTFS-----NVDLITEPIIS----- | 644 |
| Db | 62 | -----AAQFYDRLLDPLDPTYRMLPELAESWEVLIDNGATYRFHR-RDYPFQKTDWFPTP 115 | Qy | 192 | ALMLRIVHLVRD-----PRAVRSREQTAKALARDNGIVLGT-----NGTVVE 234 | | |
| Qy | 96 | ----IRSVFCLCDMDVFDAYLPWRRNIS-----DLFOQAVASRALCSPPVCEAFAARGNIS 144 | Db | 645 | -RFDILCVLVRDTPVQDEMELARPVGSHVRHPSNSKEEGLANGAEEAPAMPNTYGVF 702 | | |
| Db | 116 | RKMNADDVVTFFQRIFDRNRPW-HVNNGSNFPYDSLQPA-----DNVK 158 | Qy | 235 | ADPRLRVVNECRSHVRIAEALHKKPFPFLQDRYRLR-YEDLARDPLTVIRELYAFTGL 293 | | |
| Qy | 145 | SEEVKPLCCTRPFQELAQECASSSYSHVVLKEVRPENLQVYLPLSDPALNLRIVHLVRDP 204 | Db | 703 | PLQ-----EVLKKVYIYAKERVH-PKLNQMDODKAVKYSYDURKESENAT----- | 746 | |
| Db | 159 | S---VRKLDDNFTIVEFRLAQPPASFLWHLATHIASMSEAYRKLEKD---RQEQLDQP 212 | Qy | 294 | GIFTPLQLTWIIHNITHGGSGPGA-----RREAFKTSRSDALSVSQAWRHTLPFAKIRVY 345 | | |
| Qy | 205 | -----RAVLRSRREQ-----TAXLARDONGIVLGTNGT-----W 232 | Db | 747 | GSPITVHIESMSHGCGSPRAHPSAGLCDRERRQHGHPRDA--GEIHLRHT---EVQRH 799 | | |
| Db | 213 | VGTGPYQLSSYRAGOFIRLQRHDDEFWRGKELMPQVvvdLGGGTYGRLSKLTLGCDVLAW 272 | Qy | 346 | QUELGAGALQLLGGRYRSVHBL 365 | | |
| Qy | 233 | VEA-----DPRLRVUNEVCRSHVRIAEALHKKPFPFLQD--RYRLVRYEDLARDPL 281 | Db | 800 | RSMRKTAFARYLSFRDNNEL 819 | | |
| Db | 273 | PAASQSLILRDPPLRL---TLRPGMVNVAYLAFTAQPPLNNPAVRHALA---LATINQ 325 | | | | | |
| Qy | 282 | TVIREL---AFTGIGLGLTPQLTWHN---ITHGSGPGRARREAFKTSRSDALSV----- | Db | 90862 | hypothetical protein Ecs1871 (imported) - Escherichia coli (strain O157:H7, substrain RIM | RESULT 12 | |
| Db | 326 | RLMQSIIYYGAETAASILPR-ASWAYDNZERKITE YNPAKSREQLSKSGLENLTKLWTP 383 | C;Species: | E. coli | | | |
| Qy | 330 | --SQAWRHTLPFAKIRVQFLCGGALQIQLGYSRTHSELFCR-----DLSDLPLLPRGM 380 | C;Date: | 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004 | | | |
| Db | 384 | TRSQAWNPS PL---KTAFLIQADMAQYCVKVVIVPVERCFQEARLMMDSHDTLS--- 435 | C;Accession: | G90862 | | | |
| Qy | 381 | DSFRKASSTEXOPES 395 | A;Reference number: | A99629; PMID:21156231; PMID:11258796 | | | |
| Db | 436 | ---GMA7DS-NDPDS 446 | A;Accession: | G90862 | | | |
| | | | A;Status: | preliminary | | | |
| | | | A;Residues: | 1-547 <HAY> | | | |
| | | | A;Cross-references: | UNIPROT:Q8X7F3 ; GB:BA000007 ; PIDN:BA35294_1; PID:g13361336; GSPDB:GN | | | |
| | | | A;Experimental source: | strain O157:H7, substrain RIMD 0509952 | | | |
| | | | C;Genetics: | | | | |
| | | | A;Gene: | Ecs1871 | | | |
| | | | C;Superfamily: | di peptide transport protein | | | |
| | | | Query Match | 4.5% | Score 93; DB 2; Length 547; | | |
| | | | Best Local Similarity | 22.0% | Pred. No. 5.3; | | |
| | | | Matches | 109; | Mismatches 53; Conservative 53; Mismatches 171; | | |
| | | | A;Title: | Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genon | | | |
| | | | A;Reference number: | A99629; PMID:21156231; PMID:11258796 | | | |
| | | | A;Accession: | G90862 | | | |
| | | | A;Molecule type: | DNA | | | |
| | | | A;Status: | preliminary | | | |
| | | | A;Residues: | 1-547 <HAY> | | | |
| | | | A;Cross-references: | UNIPROT:Q8X7F3 ; GB:BA000007 ; PIDN:BA35294_1; PID:g13361336; GSPDB:GN | | | |
| | | | A;Experimental source: | strain O157:H7, substrain RIMD 0509952 | | | |
| | | | C;Genetics: | | | | |
| | | | A;Gene: | Ecs1871 | | | |
| | | | C;Superfamily: | di peptide transport protein | | | |
| | | | Query Match | 4.5% | Score 93; DB 2; Length 547; | | |
| | | | Best Local Similarity | 22.0% | Pred. No. 5.3; | | |
| | | | Matches | 109; | Mismatches 53; Conservative 53; Mismatches 171; | | |
| | | | A;Title: | Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genon | | | |
| | | | A;Reference number: | A99629; PMID:21156231; PMID:11258796 | | | |
| | | | A;Accession: | G90862 | | | |
| | | | A;Molecule type: | RNA | | | |
| | | | A;Status: | nucleic acid sequence not shown | | | |
| | | | A;Residues: | 1-892 <TOD> | | | |
| | | | A;Cross-references: | UNIPROT:P49736; EMBL:X67334; NID:9468703; PIDN:CAA47749_1; PID:94687 | | | |
| | | | A;Map position: | 3q21-3q21 | | | |
| | | | C;Complex: | The predominant form is a heterohexamer of MCM2, MCM3, MCM4, MCM5, and MCM6, and is one of several proteins that must be bound phosphorylated and dissociate from the chromatin. | | | |
| | | | C;Function: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | cdcl19 | | | |
| | | | A;Cross-references: | GDB:224876; OMIM:116945 | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches | 68; | Conservative 54; Mismatches 119; Indels 79; Gaps 15; | | |
| | | | A;Gene: | | | | |
| | | | A;Cross-references: | | | | |
| | | | A;Map position: | | | | |
| | | | C;Complex: | | | | |
| | | | A;Description: | part of the replication licensing system that permits DNA replication to proceed and of the replication licensing factor MCM2; MCM homology | | | |
| | | | C;Keywords: | cell cycle control; DNA replication initiation; heterohexameric phosphoprotein; MCM homology <MCM> F;318-344/Region: zinc finger CCCC motif | | | |
| | | | Query Match | 4.6% | Score 94.5; DB 1; Length 832; | | |
| | | | Best Local Similarity | 21.2% | Pred. No. 7.4; | | |
| | | | Matches</ | | | | |

| | | | |
|---|--|-----|--|
| | | | C.Species: Zea mays (maize) C.Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004 C.Accession: T03412 R.Pack, N.C. submitted to the EMBL Data Library, September 1994 A.Reference number: 214939 A.Accession: T03412 A.Status: Preliminary; translated from GB/EMBL/DDBJ A.Molecule type: mRNA A.Residues: 1-559 <PPE> A.Cross-references: UNIPROT:P4081; EMBL:L35914; NID:953264; PMID:AAB04118.1; PID:95326; |
| Db | 384 TRSQAWNPS-PL---KTAELIQADMAQVGVKVVIPVEGGRFQEARLMDMSHDLTLS--- | 435 | |
| Qy | 381 DSFKWASSTEKQPES 395 | | |
| Db | 436 ---GWAUDS-NDPDS 446 | | |
| RESULT 13 | | | |
| B85756 | hypothetical protein sapA [imported] - Escherichia coli (strain O157:H7, substrate EDL93 | | |
| C.Species: Escherichia coli | | | |
| C.Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004 | | | |
| R;Parra, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, I.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimambato, E.; Potamousis, K.; Apodaca, R.; Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7. | | | |
| A;Reference number: A85480; MUID:21074935; PMID:11206551 | | | |
| A;Status: preliminary | | | |
| A;Molecule type: DNA | | | |
| A;Cross-references: UNIPROT:Q8X7E3; GB:AE005174; NID:912515493; PMID:AAG56510.1; GSPPB:G | | | |
| A;Experimental source: strain O157:H7, substrate EDL93 | | | |
| C;Genetics: | | | |
| A;Gene: sapA | | | |
| C;Superfamily: dipeptide transport protein | | | |
| Query Match 4.5%; Score 93; DB 2; Length 547; | | | |
| Best Local Similarity 22.0%; Pred. No. 5.3; | | | |
| Matches 109; Conservative 53; Mismatches 171; Indels 162; Gaps 27; | | | |
| Db | 11 MLSLIMVQGIL-----VELVSRQVPS---SPAGLSEERVHVL 46 | | |
| Qy | 4 VLSLIVLVAIQLVSGQIAAPESPAPHADIRSDGTVYCVSQVNTENPSKASGLIVDTL-- | 61 | |
| Db | 47 SWSRGSSFPVGQLFQSHPDVYLMBPANEVWDTLUSQGSAPALIMAVRDL-- | 95 | |
| Qy | 62 -----AAQFYDRLLDVTDPYTRYLMPLEAESWEVDNGATYRFHLR-RDVPFQXTAWPT 115 | | |
| Qy | 96 -----IRSFLCDMDVFDAYLPWRRNIS-----DLFQWAVSRALCSPPVCEAFARGNIS 144 | | |
| Db | 116 RKNADDVVFITFORIFDRANPPW-HVNNGSNPPYFDSLQFA-----DRVKA 158 | | |
| Qy | 145 SEEYCKPLCATRPGGLAQEAFCSSYHVYLKEVRFENLQVLYPLSLSDPALNRIVRDP 204 | | |
| Db | 159 S---VRLDNHNTVERLAQDASFLWHLATHYASVMSWYARLKLEKD---RQEQLDRQ 212 | | |
| Qy | 205 -----RAVLRSRPQ-----TAKALARNDNGIVLGTNT-----W 232 | | |
| Db | 213 VGTCPYOLSEYRAQFIRLQRHDFWRGKPLMQVQVVDLGSGCTGRLSKLTGECDVLA 272 | | |
| Qy | 233 VEA -----DPLRLRVNEVCRSHVRIAALHKPPBLQD--RYRLVRYEDLARDPL 281 | | |
| Db | 273 PAASQSLSIIRRDPRLR---TLRPGMVNAYLAFTNATAKPLNNPAVRHALA---LA1NNQ 325 | | |
| Qy | 282 TVIREY---AFTCGLGLTPOLQWTIHN---IITGSGSGGARREAFKTTSRDAVL-----329 | | |
| Db | 326 RLMSOIIYGTAAETPASILPER-ASWAYDNEAKATPE-YNPKSREQLKLGLENLTKLWVP 383 | | |
| Qy | 330 --SOAWRHTLPPAKIRRVQBLCGSALQQLGYRSVHSELEQR-----DLSDLILLPRGM 380 | | |
| Db | 384 TRSQAWNPS-PL---KTAELIQADMAQVGVKVVIPVEGGRFQEARLMDMSHDLTLS--- | 435 | |
| Qy | 381 DSFKWASSTEKQPES 395 | | |
| Db | 436 ---GWAUDS-NDPDS 446 | | |
| RESULT 14 | | | |
| T03412 malate synthase (EC 4.1.3.2) - maize | | | |
| Qy | 9 TWMLSLIMVQGILVFFLVSQVPPSSPAGLGERVHVLWSSL-WRSGSSSFVGQLFSQHDFV 67 | | |

| | | | |
|----|-----|--|-----|
| Db | 3 | TVTUSIFRNFEKRLWVLQGMQTANKLGM-...-HYLPKAKFWKFMFGSGTGTGQFTPKPN-- | 56 |
| Qy | 68 | YLMBPAWHDYDTLSOGSAPALHMAVRDLIRSVFLCDMDVFDAYLPWNRNISDLFOWAVSR | 127 |
| Db | 57 | -----WHYSILAWW--PDEETARREVAES-----PIYORWTMADESY-----T | 94 |
| Qy | 128 | ALCSPPVCEFAFGNISSEEVCKPLCATRPGFLAQEAC-----SSYSHVV-- | 172 |
| Db | 95 | VLLQPTSANGKWDGKEPFPEV-KPASDVRPIAULTRAYTFWKWAERFGARAATISHMIGR | 153 |
| Qy | 173 | LKEYRF-----FNLOVLYPLISDPAINLRLITVHLVRDRAVRSRECTAKALARDNGI | 224 |
| Db | 154 | NKDUVFKIGVEPVQQVTSFSINPDAARWR----SSPRGAGGPHGEAIKAVALAENWF | 208 |
| Qy | 225 | -----VLTGTINGTWEDAPRLRVVNEVCRSH-----VRAEAALHK | 259 |
| Db | 209 | KEELYYARFQLGTIGKWEGRDP---VGZALTARPSEAKPAPAPAEPKPA | 264 |
| Qy | 260 | PPPFHQDRYRLVRYEDLARDPLTVRELXAFTGGL | 295 |
| Db | 265 | PAPVAEKPALAVENMKPABEPKPVVEAPKRPFRSRGL | 300 |

Search completed: June 23, 2005, 08:50:11
 Job time : 11.953 secs

THIS PAGE BLANK (USPTO)

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|---|
| 1 | 2057 | 100.0 | 395 | 16 US-10-697-828-7 Sequence 7, Appli |
| 2 | 2057 | 100.0 | 395 | 16 US-10-697-828-7 Sequence 8, Appli |
| 3 | 2042.5 | 99.3 | 418 | 9 US-10-841-707-8 Sequence 5, Appli |
| 4 | 1727 | 84.0 | 395 | 9 US-09-927-002-5 Sequence 3, Appli |
| 5 | 1708 | 83.0 | 395 | 9 US-09-927-002-3 Sequence 2, Appli |
| 6 | 1708 | 83.0 | 395 | 15 US-10-258-080-5 Sequence 5, Appli |
| 7 | 1708 | 83.0 | 395 | 16 US-10-648-593-159 Sequence 13, Appli |
| 8 | 1708 | 83.0 | 395 | 16 US-10-697-828-13 Sequence 4, Appli |
| 9 | 1536.5 | 74.7 | 390 | 9 US-09-927-002-4 Sequence 8, Appli |
| 10 | 1536.5 | 74.7 | 390 | 16 US-10-697-828-8 Sequence 2211, Ap |
| 11 | 1333 | 64.8 | 394 | 16 US-10-408-765A-2211 |

| | | | | | | | |
|----|-----|---|---------------------|-----------|-------------|---|------|
| Db | 61 | SQHPDVYLMEPAWHWDLTLSQSSAPALHMAVRDLIRSVFLCDMDVFDAYLPWRRNISDL | 120 | Db | 301 | TWIINITHGSGPGRAREAFKTTSRDALSYSQAWRHTLPPAKIRVQELCGGALQLLGYSR | 360 |
| Qy | 121 | FOWAVSRAICSPPVCEAARGNTISSEEVCKPLCATRPGFLAQEACSSYSHVVKERYFFN | 180 | Qy | 361 | VHSELQRDLSLDLJLPRGMDSFKWASSTEKOPES | 3 95 |
| Db | 121 | FOWAVSRAICSPPVCEAARGNTISSEEVCKPLCATRPGFLAQEACSSYSHVVKERYFFN | 180 | Db | 361 | VHSELQRDLSLDLJLPRGMDSFKWASSTEKOPES | 3 95 |
| Qy | 181 | LQVLYPLISDPAINLRVHLVDPRAVLRSRECTAKALARNDIVIQTGTVWEAPRL | 240 | | | | |
| Db | 181 | LQVLYPLISDPAINLRVHLVDPRAVLRSRECTAKALARNDIVIQTGTVWEAPRL | 240 | | | | |
| Qy | 241 | VNEVCRSHVRIAEEALKPPPLQDRYRLVRELYAFTGLTGPQLQ | 300 | | | | |
| Db | 241 | VNEVCRSHVRIAEEALKPPPLQDRYRLVRELYAFTGLTGPQLQ | 300 | | | | |
| Qy | 301 | TWIINITHGSGPGRAREAFKTTSRDALSYSQAWRHTLPEAKIRVQELCGGALQLLGYSR | 360 | | | | |
| Db | 301 | TWIINITHGSGPGRAREAFKTTSRDALSYSQAWRHTLPEAKIRVQELCGGALQLLGYSR | 360 | | | | |
| Qy | 361 | VHSELQRDLSLDLJLPRGMDSFKWASSTEKOPES | 3 95 | | | | |
| Db | 361 | VHSELQRDLSLDLJLPRGMDSFKWASSTEKOPES | 3 95 | | | | |
| | | RESULT 3 | US-09-927-602-5 | | | | |
| | | Sequence 5, Application US/0927602 | | | | | |
| | | Patent No. US20020061562A1 | | | | | |
| | | GENERAL INFORMATION: | | | | | |
| | | APPLICANT: Fukuda, Michiko N. | | | | | |
| | | TITLE OF INVENTION: Methods of Treating Macular Corneal Dystrophy | | | | | |
| | | FILE REFERENCE: P-LJ 4852 | | | | | |
| | | CURRENT APPLICATION NUMBER: US/09/927,602 | | | | | |
| | | CURRENT FILING DATE: 2001-08-09 | | | | | |
| | | PRIOR APPLICATION NUMBER: US 09/638,211 | | | | | |
| | | PRIOR FILING DATE: 2000-08-11 | | | | | |
| | | NUMBER OF SEQ ID NOS: 38 | | | | | |
| | | SOFTWARE: FastSEQ for Windows Version 4.0 | | | | | |
| | | SEQ ID NO 5 | | | | | |
| | | LENGTH: 418 | | | | | |
| | | ORGANISM: Mus musculus | | | | | |
| | | TYPE: PRT | | | | | |
| | | US-09-927-602-5 | | | | | |
| | | Query Match 99.3% | Score 2042.5; | DB 9; | Length 418; | | |
| | | Best Local Similarity 99.7%; | Pred. No. 2.8e-194; | | | | |
| | | Matches 394; Conservative 0; | Mismatches 0; | Indels 1; | Gaps 1; | | |
| | | Query 1 MRLPRESSTMISLIMVQTGIVLFLVSYQVSSPAGLGERVHVLVLSNRSGSSVGGOLF 60 | | | | | |
| | | Db 25 MRLPRESSTMISLIMVQTGIVLFLVSYQVSSPAGLGERVHVLVLSNRSGSSVGGOLF 84 | | | | | |
| | | Qy 61 SQHPDVYLMEPAMHWTUQSAPALHMAVRDLIRSVELCDMDVFDAYLPWRRNISDL 120 | | | | | |
| | | Db 85 SQHPDVYLMEPAMHWTUQSAPALHMAVRDLIRSVELCDMDVFDAYLPWRRNISDL 144 | | | | | |
| | | Qy 121 FQAWYRALCSPPVCEAFARGNISSEEVCKPLCATRPGFLAQEACSSYSHVVKERVFFN 180 | | | | | |
| | | Db 145 FQAWYRALCSPPVCEAFARGNISSEEVCKPLCATRPGFLAQEACSSYSHVVKERVFFN 204 | | | | | |
| | | Qy 181 LQVLYPLISDPAINLRVHLVDRPAVLISREOTAKALARNDIVIQTGTVWEADPRLR 240 | | | | | |
| | | Db 205 LQVLYPLISDPAINLRVHLVDRPAVLISREOTAKALARNDIVIQTGTVWEADPRLR 264 | | | | | |
| | | Qy 241 VVNEVCRSHVRIAEEALKPPFLDRYRLVRYEELDLTIRELYAFTGLTGPQLQ 300 | | | | | |
| | | Db 265 VVNEVCRSHVRIAEEALKPPFLDRYRLVRYEELDLTIRELYAFTGLTGPQLQ 323 | | | | | |
| | | Qy 301 TWIINITHGSGPGRAREAFKTTSRDALSYSQAWRHTLPEAKIRVQELCGGALQLLGYSR | | | | | |
| | | Db 324 TWIINITHGSGPGRAREAFKTTSRDALSYSQAWRHTLPEAKIRVQELCGGALQLLGYSR | | | | | |
| | | RESULT 4 | US-09-927-602-3 | | | | |
| | | Sequence 3, Application US/0927602 | | | | | |
| | | Patent No. US20020061562A1 | | | | | |
| | | GENERAL INFORMATION: | | | | | |
| | | APPLICANT: Fukuda, Michiko N. | | | | | |
| | | TITLE OF INVENTION: Methods of Treating Macular Corneal Dystrophy | | | | | |
| | | FILE REFERENCE: P-LJ 4852 | | | | | |
| | | CURRENT APPLICATION NUMBER: US/09/927,602 | | | | | |

CURRENT FILING DATE: 2001-08-09
 PRIORITY FILING DATE: 2000-08-11
 NUMBER OF SEQ ID NOS: 38
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 3
 LENGTH: 395
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic construct
 NAME/KEY: VARIANT
 LOCATION: (1) ..(395)
 OTHER INFORMATION: xaa = any amino acid

Query Match Score 1727; DB 9; Length 395;
 Best Local Similarity 84.7%; Pred. No. 6..8e-163;
 Matches 333; Conservative 16; Mismatches 44; Indels 0; Gaps 0;

Qy 1 MRLPFFSSTVMSLIMVOTGILVFLVSRQVPPSPAGLGERVHVLSSRSGSSFVGOLF 60
 Db 1 MWLPRVSSSTAVTALLQTLQTLFLYSPGPSSPAGEARVHVLSSRSGSSFVGOLF 60
 Qy 61 SQHPDVYLMEPAHWDTLSQGSAPALHMAVRDLIRSYFLCDMDVFDAYLPWRNNSDL 120
 Db 61 SQHPDVYLMEPAHWDTLSQGSAPALHMAVRDLIRSYFLCDMDVFDAYLPWRNNSDL 120
 Qy 121 FQAVSRALCSPPVCEAFARGNSSEEVCKPLCATRPGLAQCSSYSHVVLKEVRFN 180
 Db 121 FQAVSRALCSPPVCEAFARGNSSEEVCKPLCATRPGLAQCSSYSHVVLKEVRFN 180
 Qy 181 LQVLYPLISDPALNRLVHLVRDPAVRSRECTAKALARNDIVLGTINGTWADPRLR 240
 Db 181 LQVLYPLISDPALNRLVHLVRDPAVRSRECTAKALARNDIVLGTINGTWADPGLR 240
 Qy 181 FQAVSRALCSPPVCEAFARGNSSEEVCKPLCATRPGLAQCSSYSHVVLKEVRFN 180
 Db 181 FQAVSRALCSPPVCEAFARGNSSEEVCKPLCATRPGLAQCSSYSHVVLKEVRFN 180
 Qy 241 VNEVCRSHVRIAELAHKPPFLQDRYFLVRELYAFTRGGLQLLGYS 360
 Db 301 AWIHNITHGGPGGAREAFFKTSRDALSVQAWRHTLPFAKIRYQELCGALQLGTRP 360
 Qy 241 VNEVCRSHVRIAELAHKPPFLQDRYFLVRELYAFTRGGLQLLGYS 360
 Db 301 AWIHNITHGGPGGAREAFFKTSRDALSVQAWRHTLPFAKIRYQELCGALQLGTRP 360
 Qy 361 VHSELEQDLSLDLPLPGMDSFKWASSTEKOP 395
 Db 361 VYSEDEQRDLALDVLVRGLNGFTWAASSASHPRN 395

RESULT 6
 US-10-258-080-5
 Sequence 5, Application US/10258080
 Publication No. US20040039125A1
 GENERAL INFORMATION:
 APPLICANT: Incyte Genomics, Inc.
 ATTORNEY/AGENT: POLICK, Jennifer L.
 HAFALIA, April J.A.
 APPLICANT: BURFORD, Neil
 RING, Huijun Z.
 APPLICANT: LAL, Preeti G.
 APPLICANT: TRIBOURBY, Catherine M.
 APPLICANT: YAO, Monique G.
 APPLICANT: YUE, Henry
 APPLICANT: TANG, Y. Tom
 APPLICANT: ARIZU, Chandra S.
 APPLICANT: DAS, Debopriya
 APPLICANT: SANJAYALA, Madhusudan M.
 APPLICANT: GANDHI, Ameena R.
 APPLICANT: REDDY, Roopa M.
 APPLICANT: KHAN, Farrah A.
 APPLICANT: BAUGH, Mariah R.
 APPLICANT: RAMKUMAR, Jayalakshmi
 APPLICANT: GRIFFIN, Jennifer A.
 APPLICANT: AU-YOUNG, Janie K.
 TITLE OF INVENTION: DRUG METABOLIZING ENZYMES

FILE REFERENCE: PI-0070 USN
 CURRENT APPLICATION NUMBER: US/10/258, 080
 CURRENT FILING DATE: 2002-10-15
 PRIOR APPLICATION NUMBER: US 60/203, 509
 PRIOR FILING DATE: 2000-05-11
 PRIOR APPLICATION NUMBER: US 60/202, 234
 PRIOR FILING DATE: 2000-05-05
 PRIOR APPLICATION NUMBER: US 60/200, 185
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: US 60/198, 403
 PRIOR FILING DATE: 2000-04-19
 PRIOR APPLICATION NUMBER: PCT/US01/11869
 PRIOR FILING DATE: 2001-04-12
 PRIOR APPLICATION NUMBER: US 60/197, 590
 PRIOR FILING DATE: 2000-04-13
 NUMBER OF SEQ ID NOS: 20

Query Match Score 1708; DB 9; Length 395;
 Best Local Similarity 83.0%; Pred. No. 5..3e-161;

Qy 1 Sequence 2, Application US/09921602
 PATENT NO.: US20020061562A1
 GENERAL INFORMATION:
 APPLICANT: Fukuda, Michiko N.
 APPLICANT: Akana, Tomoya O.
 TITLE OF INVENTION: Methods of Treating Macular Corneal Dystrophy
 FILE REFERENCE: P-LJ 4852
 CURRENT APPLICATION NUMBER: US/09/927, 602
 CURRENT FILING DATE: 2001-08-09
 PRIOR APPLICATION NUMBER: US 09/638, 211
 PRIOR FILING DATE: 2000-08-11
 NUMBER OF SEQ ID NOS: 38
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 2
 LENGTH: 395
 TYPE: PRT
 ORGANISM: Homo Sapien

Qy 1 Query Match Score 1708; DB 9; Length 395;
 Best Local Similarity 83.0%; Pred. No. 5..3e-161;

Qy 1 US-09-927-602-2

SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY:
; OTHER INFORMATION: Incyte ID No. US20040029125A1 7472777CD1

US-10-258-080-5

Query Match 83.0%; Score 1708; DB 15; Length 395;
Best Local Similarity 83.0%; Pred. No. 5.3e-161;
Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MRLPREFSTVMLSLIMVQTGILFLVLSRQVSSPAIGLGERHVLVLSWRSSESSFVGOLF 60
Db 1 MWLPRVSSTAVTALLAQTLFLFLVSRGPSSPAGGEARHVYLSSWRSGSSFVGOLF 60
Qy 61 QHPDPVYLMPEAHWDTLISQSAPALHMAVDLIRSVCMDVDAYLPWRRNISDL 1.20
Db 61 NQHPDVYLMPEAHWWTLLSQSAAATHMABDVLVRSEFLCDMDVDAYLPWRRNISDL
Qy 121 FOWAVSRLCSPPVCEAFARGNISSEEVCKPLCATRPGLAQFACSSYSHVILKEVRFN 180
Db 121 FOWAVSRLCSPPCSPACSFPRGASSEANCKPLCARQSFTLAREACSYSHVILKEVRFN 180

RESULT 8
US-10-697-828-13
; Sequence 13, Application US/10697828
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven
; APPLICANT: Lee, Jin Kyu
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfo transferases GST-4alpha, GST-4beta, & GST-6
; CURRENT APPLICATION NUMBER: US/10/697,828
; CURRENT FILING DATE: 2003-10-29
; PRIORITY NUMBER: 09/593,828
; PRIORITY FILING DATE: 2000-06-13
; PRIORITY NUMBER: 60/144,694
; PRIORITY FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 395
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-697-828-13

Query Match 83.0%; Score 1708; DB 16; Length 395;
Best Local Similarity 83.0%; Pred. No. 5.3e-161;
Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MRLPREFSTVMLSLIMVQTGILFLVLSRQVSSPAGLGERHVLVLSWRSSESSFVGOLF 60
Db 1 MWLPRVSSTAVTALLAQTLFLFLVSRGPSSPAGGEARHVYLSSWRSGSSFVGOLF 60

Qy 121 FOWAVSRLCSPPVCEAFARGNISSEEVCKPLCATRPGLAQFACSSYSHVILKEVRFN 180
Db 121 FOWAVSRLCSPPCSPACSFPRGASSEANCKPLCARQSFTLAREACSYSHVILKEVRFN 180

Query Match 83.0%; Score 1708; DB 16; Length 395;
Best Local Similarity 83.0%; Pred. No. 5.3e-161;
Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MRLPREFSTVMLSLIMVQTGILFLVLSRQVSSPAIGLGERHVLVLSWRSSESSFVGOLF 60
Db 1 MWLPRVSSTAVTALLAQTLFLFLVSRGPSSPAGGEARHVYLSSWRSGSSFVGOLF 60

Qy 181 QVNEVCRSHVRLIAEAALHKKPPLQDRYLVLDLPRDPLVIRELYAFTGLGLTFQLQ 300
Db 181 LQVLYPLISDPALNLRVHLVYRDRPLRSREOTAKALARDNGIVLTGNTWVEADPLR 240
Qy 241 VVNEVCRSHVRLIAEAALHKKPPLQDRYLVLDLPRDPLVIRELYAFTGLGLTFQLQ 300
Db 241 VVREVCRSHVRLIAEAATLKPKPLFRGRYLVRFEDAREPLAEIRAYAFTGLGLTFQLQ 300

Qy 241 VVREVCRSHVRLIAEAALHKKPPLQDRYLVLDLPRDPLVIRELYAFTGLGLTFQLQ 300
Db 241 VVREVCRSHVRLIAEAATLKPKPLFRGRYLVRFEDAREPLAEIRAYAFTGLGLTFQLQ 240

Qy 241 VVREVCRSHVRLIAEAALHKKPPLQDRYLVLDLPRDPLVIRELYAFTGLGLTFQLQ 300
Db 241 VVREVCRSHVRLIAEAATLKPKPLFRGRYLVRFEDAREPLAEIRAYAFTGLGLTFQLQ 300

Qy 301 TWIHNTHGSGPGRFAKTTSDALSVS0A9RHTLPPAKIRRVQELCGGALQLGYS 360
Db 301 AWIHNTHGSGPGRFAKTTSDALSVS0A9RHTLPPAKIRRVQELCGGALQLGYS 360

Qy 361 VHSELEQDLSLILLPRGMSDFKWASSTEKEOPES 395
Db 361 VYSEDEQRNLALDLVLRGLNGFTWASSTASHPRN 395

RESULT 9
US-10-697-828-13
; Sequence 13, Application US/10697828
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven
; APPLICANT: Lee, Jin Kyu
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfo transferases GST-4alpha, GST-4beta, & GST-6
; CURRENT APPLICATION NUMBER: US/10/697,828
; CURRENT FILING DATE: 2003-10-29
; PRIORITY NUMBER: 09/593,828
; PRIORITY FILING DATE: 2000-06-13
; PRIORITY NUMBER: 60/144,694
; PRIORITY FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 395
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-697-828-13

Query Match 83.0%; Score 1708; DB 16; Length 395;
Best Local Similarity 83.0%; Pred. No. 5.3e-161;
Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MRLPREFSTVMLSLIMVQTGILFLVLSRQVSSPAGLGERHVLVLSWRSSESSFVGOLF 60
Db 1 MWLPRVSSTAVTALLAQTLFLFLVSRGPSSPAGGEARHVYLSSWRSGSSFVGOLF 60

Qy 61 QHPDPVYLMPEAHWDTLISQSAPALHMAVDLIRSVCMDVDAYLPWRRNISDL 1.20
Db 61 NQHPDVYLMPEAHWWTLLSQSAAATHMABDVLVRSEFLCDMDVDAYLPWRRNISDL 1.20

Qy 121 FOWAVSRLCSPPVCEAFARGNISSEEVCKPLCATRPGLAQFACSSYSHVILKEVRFN 180
Db 121 FOWAVSRLCSPPCSPACSFPRGASSEANCKPLCARQSFTLAREACSYSHVILKEVRFN 180

Query Match 83.0%; Score 1708; DB 16; Length 395;
Best Local Similarity 83.0%; Pred. No. 5.3e-161;
Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0;

Qy 1 MRLPREFSTVMLSLIMVQTGILFLVLSRQVSSPAIGLGERHVLVLSWRSSESSFVGOLF 60
Db 1 MWLPRVSSTAVTALLAQTLFLFLVSRGPSSPAGGEARHVYLSSWRSGSSFVGOLF 60

Qy 181 QVNEVCRSHVRLIAEAALHKKPPLQDRYLVLDLPRDPLVIRELYAFTGLGLTFQLQ 300
Db 181 VVREVCRSHVRLIAEAATLKPKPLFRGRYLVRFEDAREPLAEIRAYAFTGLGLTFQLQ 300

Qy 241 VVREVCRSHVRLIAEAALHKKPPLQDRYLVLDLPRDPLVIRELYAFTGLGLTFQLQ 300
Db 241 VVREVCRSHVRLIAEAATLKPKPLFRGRYLVRFEDAREPLAEIRAYAFTGLGLTFQLQ 300

RESULT 9
 Qy 301 TWINNITHGSGP GARREAFKTTSDALSVSQAWRHTLPPAKIRVQELGGALQLLGYRS 360
 Db 301 AWINNITHGSGP GARREAFKTTSDALSVSQAWRHTLPPAKIRVQELGGALQLLGYRS 360
 Qy 361 VHSLEQRLDLSLDPGMDSFKAWSSTEKQPES 395
 Db 361 VYSEDBQRNLALDLVLPGRINGFTWAStASHPRN 395
 SEQ ID NO: 4
 LENGTH: 390
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-697-828-8

RESULT 9
 Sequence 4, Application US/09927602
 Patent No. US2002006162A1
 GENERAL INFORMATION:
 APPLICANT: Fukuda, Michiko N.
 TITLE OF INVENTION: Methods of Treating Macular Corneal
 CURRENT FILING DATE: 2001-08-09
 PRIOR APPLICATION NUMBER: US/09/927,602
 PRIORITY NUMBER: 09/638,211
 PRIOR FILING DATE: 2000-08-11
 NUMBER OF SEQ ID NOS: 38
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 4
 LENGTH: 390
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-927-602-4

Query Match 74.7%; Score 1536.5; DB 9; Length 390;
 Best Local Similarity 76.0%; Pred. No. 6.3e-144; Indels 1; Gaps 1;
 Matches 295; Conservative 30; Mismatches 62; Indels 1; Gaps 1;
 Qy 1 MRLPRESSEDPTVMSLIMMVQTG-ILVFLVSRQVPSSPAAGLGERVHVLSSWRSGSSFVGOL 59
 Db 1 MWLPRFESSPTVMSLIMMVQTG-ILVFLVSRQVPSSPAAGLGERVHVLSSWRSGSSFVGOL 60
 Qy 60 FSQHPDVFLYMEPAWHDWLTSQGSSAPALHMAVDRDLTSVPLCDMDVFDAYLPWRNISD 119
 Db 61 FSQHPDVFLYMEPAWHDWLTSQGSSAPALHMAVDRDLTSVPLCDMDVFDAYMPOSRNLSA 120
 Qy 120 LFQWAVSRAALCSPPACSAFPARGNTSSEVKPGLAQASCSYSHVILKEVRFF 179
 Db 121 FFWNTSRAALCSPPACSAFPARGNTSKQDCVKTLCTRQPPLSLAREACSYSHVILKEVRFF 180
 Qy 180 NLQVLYPLISDPALNLRIVLHYRDPRAVLRSREOTAKALARDNGVILCTINGTWAEADPRL 239
 Db 181 NLQVLYPLISDPALNLRIVLHYRDPRAVLRSREAGPPIALARDNGVILCTINGKWEADPRL 240
 Qy 240 RVVNEVCRSHVRFAEAUHKKPPEFLQDPRYLVRYEDLARDPLTVIRELYAFTGLGLTPOL 299
 Db 241 RLIREVCRSHVRFAEAUHKKPPEFLQDPRYLVRYEDLARDPLTVIRELYAFTGLGLTPOL 300
 Qy 300 QTWHNITHGSGPAAEFGTTSRDLAUSVQWRHTLPPFAKIRRVQELGGALQLLGYR 359
 Db 301 EAWHNITHGSGIGKPIAEHTTSRNARNSQAWRHALPFKILRQEVCAGALQLLGYR 360
 Qy 360 SVHSELEQRLDLSLDPGMDSFKWAS 387
 Db 361 PVTSADQRLDLSLDPGMDSFKWAS 388

RESULT 11
 US-10-408-765A-2211
 Sequence 2211, Application US/10408765A
 PUBLICATION NO. US20040101874A1
 GENERAL INFORMATION:
 APPLICANT: Ghosh, Soumitra S.
 APPLICANT: Fahy, Eoin D.
 APPLICANT: Zhang, Bing
 APPLICANT: Gibson, Bradford W.
 APPLICANT: Taylor, Steven W.
 APPLICANT: Glenn, Gary N.
 APPLICANT: Warnock, Dale E.
 TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 FILE REFERENCE: 660088_465
 CURRENT APPLICATION NUMBER: US/10/408765A
 CURRENT FILING DATE: 2003-04-04
 NUMBER OF SEQ ID NOS: 3077
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 2211
 LENGTH: 394
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-408-765A-2211
 RESULT 10
 US-10-697-828-8
 Sequence 8, Application US/10697828
 Publication No. US20040185546A1
 GENERAL INFORMATION:
 APPLICANT: Rosen, Steven
 APPLICANT: Lee, Jin Kyu
 APPLICANT: Hemmerich, Stefan
 TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-6

Query Match 64.8%; Score 1333; DB 16; Length 394;
 Best Local Similarity 78.1%; Pred. No. 1.2e-123;
 Matches 257; Conservative 19; Mismatches 47; Indels 6; Gaps 1;

Db 239 EDQPYYMQVICQSLEYK-TIOSLPKALOERYLLVRYEDLARAPVACTSRMYEFFVGL 297
 Qy 295 LTPQLOQTWHNITHGSGPGRAREAFKTSRDSLVSQAWRHTLPFAKIRYQELCGGALQ 354
 Db 298 FLPHQQTWNITRGMG - DHAFTINARDALNYSQAWRNSLPPYEVSRLOQACGDMN 355
 Qy 355 LLGYRSVHSELEQRDLSDL 375
 Db 356 LLGYRHVRSEOEQRNLSDL 376

RESULT 13
 US-10-007-262-1
 ; Sequence 1, Application US/10007262
 ; Publication No. US200216478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bistrop, Annette
 ; APPLICANT: Rosen, Steven D.
 ; APPLICANT: Tangemann, Kirsten
 ; APPLICANT: Hemmerich, Stefan
 ; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
 ; FILE REFERENCE: 6510-107CIP
 ; CURRENT APPLICATION NUMBER: US/10/007,262
 ; CURRENT FILING DATE: 2001-11-08
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
 ; NUMBER OF SEQ ID NOS: 8
 ; SEQ ID NO: 1
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: H. sapiens
 ; US-10-007-262-1

Query Match 49.5%; Score 1017.5; DB 13; Length 386;
 Best Local Similarity 56.4%; Pred. No. 3e-92;
 Matches 215; Conservative 43; Mismatches 112; Indels 11; Gaps 7;

Db 1 MLLPK-KMPLLFLPVSMATIALFPHMYSNINSSLMAOPERMHFLVLSWRSSSFV 58
 Qy 1 MRLPRPSTVMSLINVQTGILVF---LVSROVPS-SPAGLGERVHVLSSWRSGSSFV 56
 Db 1 MLLPK-KMPLLFLPVSMATIALFPHMYSNINSSLMAOPERMHFLVLSWRSSSFV 58
 Qy 57 GOLFQHPDVFLMEPAWHYDTLSQGQSAPALHMAVDRDLIRSVELCDMDYEDAYL-PWRR 115
 Db 59 GOLFQHPDVFLMEPAWHYNTFKOSTAWMLHMAVDRDLTRAVFLCDMSVDAYNPGP 118
 Qy 116 NISDLFQWASBALCSPPVCBAFARGNISSEBVCKPLCATRPFLQIAEACSSYHVVLKE 175
 Db 119 QSSLFQWENSRALCSAPACDTIPRAHCRLCSQQPEVVKACRSYHVVLKE 178
 Qy 176 VRFENIQVYPLSLPDALNIRVHLYRDPRAVLRSREOTAKALARDNGIVLGNTGTWV-E 234
 Db 177 VRFENIQVYPLSLPDALNIRVHLYRDPRAVLRSREOTAKALARDNGIVLGNTGTWV-E 234
 Qy 235 ADPLRUVVNECRSHVRIAEEAHHKPPFLQDRYRLVRYEDLARDPLTVIRELYAFTGLG 294
 Db 239 EDQPYYMQVICQSLEYK-TIOSLPKALOERYLLVRYEDLARAPVACTSRMYEFFVGL 297

Query Match 49.5%; Score 1017.5; DB 9; Length 386;
 Best Local Similarity 56.4%; Pred. No. 3e-92;
 Matches 215; Conservative 43; Mismatches 112; Indels 11; Gaps 7;

Db 1 MLLPK-KMPLLFLVSMATIALFPHMYSNINSSLMAOPERMHFLVLSWRSGSSFV 58
 Qy 57 GOLFQHPDVFLMEPAWHYDTLSQGQSAPALHMAVDRDLIRSVELCDMDYEDAYL-PWRR 115
 Db 59 GOLFQHPDVFLMEPAWHYNTFKOSTAWMLHMAVDRDLTRAVFLCDMSVDAYNPGP 118
 Qy 116 NISDLFQWASBALCSPPVCBAFARGNISSEBVCKPLCATRPFLQIAEACSSYHVVLKE 175
 Db 119 QSSLFQWENSRALCSAPACDTIPRAHCRLCSQQPEVVKACRSYHVVLKE 178
 Qy 176 VRFENIQVYPLSLPDALNIRVHLYRDPRAVLRSREOTAKALARDNGIVLGNTGTWV-E 234
 Db 177 VRFENIQVYPLSLPDALNIRVHLYRDPRAVLRSREOTAKALARDNGIVLGNTGTWV-E 234
 Qy 235 ADPLRUVVNECRSHVRIAEEAHHKPPFLQDRYRLVRYEDLARDPLTVIRELYAFTGLG 294
 Db 239 EDQPYYMQVICQSLEYK-TIOSLPKALOERYLLVRYEDLARAPVACTSRMYEFFVGL 297

RESULT 14
 US-10-841-707-6
 ; Sequence 6, Application US/10841707
 ; Publication No. US20040202649A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Minoru

APPLICANT: Yeh, Jiunn-Chern
 APPLICANT: Hirao, Nobuyoshi
 TITLE OF INVENTION: Identification of the Meca-79 Antigen
 TITLE OF INVENTION: and Related Methods of Treating L-Selectin-Mediated
 FILE REFERENCE: P-LJ 4149
 CURRENT APPLICATION NUMBER: US/10/841,707
 PRIOR APPLICATION NUMBER: 2004-05-06
 PRIOR FILING DATE: 2000-05-11
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 6
 LENGTH: 380
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-841-707-6

Query Match 49.4%; Score 1015.5; DB 16; Length 380;
 Best Local Similarity 57.1%; Pred. No. 4.6e-92; Indels 9; Gaps 6;

Matches 212; Conservative 42; Mismatches 108; Indels 9; Gaps 6;

Qy 11 MLSLMVQTMLVF--LYSRQVPS-SPAGLGERVHVLVLSMRSGSSSTVGOLESQHPPDV 66
 Db 3 LLLFLVSOMLAIIALPFHMISHNISLSMVKQPERMHLVLSMRSGSSSEVGQFGQHPVV 62
 Qy 67 FYLMEPAHWYHWDTLSQGSAPALHNAVRDLRSVPLCDMDVFDAYL-PWFRNISDLPQAV 125
 Db 63 FYLMEPAHWYHWTMFKOSTAMLHNAVRDLTRAVFLCDMSVDAVMEPGRQSSLFQHEN 122
 Qy 126 SRALCSPPVCEAFARNNISSEEVCKPLCATRTPFGLAQEAQSSYSHVILKEVRFENLQLVY 185
 Db 123 SRALCSAPADLCPDDEIIPRAHCRLLCSCQQPEVVEKARCSYSHVILKEVRFENLQLY 182
 Qy 186 PLISDPAHLRIVHLYDPAVLRSLREQTAKALARNDGTVLGTNGTW-BADPLRVNE 244
 Db 183 PLLKDPDSLNLHVLVLDPAVFRSERTKGDLMDSRIVMGQEOKLKKEDQPYVYNNQV 242
 Qy 245 VCRSHVRIAAALHKPPPFQDRYTLVRYEDLARDPLTVIRELIAFTGLGLTPOLQTMH 304
 Db 243 ICQSQEUYK-TIQSLPKALQERYLVLRTEDLARPAVQTSRMYEFVGLBLPHLQTWVH 301
 Qy 305 NITHGSGPGARREAFKTTSDALSYSQANHTLPFAKIRVQELCGGALQOLLGYRSVHSE 364
 Db 302 NITRGKGMG--DHAPHTNADALNYSQAWNSLPSBKVSRSLPQKACGDANLLGRHVRE 359
 Qy 365 LEORDLSDLL 375
 Db 360 QEQRNLLDLL 370

RESULT 15
 US-10-427-631-11
 Sequence 11, Application US/10427631
 Publication No. US20030175923A1
 GENERAL INFORMATION:
 APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
 APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
 APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
 APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
 APPLICANT: AZIMZAI, Valda
 TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 FILE REFERENCE: PF-0522-1 DIV
 CURRENT APPLICATION NUMBER: US/10/427,631
 CURRENT FILING DATE: 2003-04-29
 PRIOR APPLICATION NUMBER: US 09/786,240
 PRIOR FILING DATE: 2002-03-12
 PRIOR APPLICATION NUMBER: PCT/US99/20989
 PRIOR FILING DATE: 1999-09-09
 PRIOR APPLICATION NUMBER: US 60/172,220
 PRIOR APPLICATION NUMBER: US 60/155,248
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: US 60/155,248
 PRIOR FILING DATE: 1998-11-04

Search completed: June 23, 2005, 09:27:52
 Job time : 51.935 secs

THIS PAGE BLANK (USPTO)

| Result | No. | Score | Query | Match | Length | DB | ID | Description |
|--------|--------|-------|-------|-------|--------|---------------------|----|---------------------|
| 1 | 2051 | 100.0 | | 390 | 9 | US-09-927-602-4 | | Sequence 4, Appli |
| 2 | 2051 | 100.0 | | 390 | 16 | US-10-697-828-8 | | Sequence 8, Appli |
| 3 | 1758.5 | 85.7 | | 395 | 9 | US-09-927-602-3 | | Sequence 3, Appli |
| 4 | 1729.5 | 84.3 | | 395 | 9 | US-09-927-602-2 | | Sequence 2, Appli |
| 5 | 1729.5 | 84.3 | | 395 | 15 | US-10-258-080-5 | | Sequence 5, Appli |
| 6 | 1729.5 | 84.3 | | 395 | 16 | US-10-648-593-159 | | Sequence 159, Appli |
| 7 | 1729.5 | 84.3 | | 395 | 16 | US-10-697-828-13 | | Sequence 13, Appli |
| 8 | 1536.5 | 74.9 | | 395 | 16 | US-10-697-828-7 | | Sequence 7, Appli |
| 9 | 1536.5 | 74.9 | | 395 | 16 | US-10-841-707-8 | | Sequence 8, Appli |
| 10 | 1523 | 74.3 | | 418 | 9 | US-09-927-602-5 | | Sequence 5, Appli |
| 11 | 1513.5 | 73.8 | | 394 | 16 | US-10-408-765A-2211 | | Sequence 2211, Ap |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query | Match | Length | DB | ID | Description |
|--------|--------|-------|-------|-------|--------|---------------------|----|---------------------|
| 1 | 2051 | 100.0 | | 390 | 9 | US-09-927-602-4 | | Sequence 4, Appli |
| 2 | 2051 | 100.0 | | 390 | 16 | US-10-697-828-8 | | Sequence 8, Appli |
| 3 | 1758.5 | 85.7 | | 395 | 9 | US-09-927-602-3 | | Sequence 3, Appli |
| 4 | 1729.5 | 84.3 | | 395 | 9 | US-09-927-602-2 | | Sequence 2, Appli |
| 5 | 1729.5 | 84.3 | | 395 | 15 | US-10-258-080-5 | | Sequence 5, Appli |
| 6 | 1729.5 | 84.3 | | 395 | 16 | US-10-648-593-159 | | Sequence 159, Appli |
| 7 | 1729.5 | 84.3 | | 395 | 16 | US-10-697-828-13 | | Sequence 13, Appli |
| 8 | 1536.5 | 74.9 | | 395 | 16 | US-10-697-828-7 | | Sequence 7, Appli |
| 9 | 1536.5 | 74.9 | | 395 | 16 | US-10-841-707-8 | | Sequence 8, Appli |
| 10 | 1523 | 74.3 | | 418 | 9 | US-09-927-602-5 | | Sequence 5, Appli |
| 11 | 1513.5 | 73.8 | | 394 | 16 | US-10-408-765A-2211 | | Sequence 2211, Ap |

Software: FastSEQ for Windows Version 4.0

SEQ ID NO: 4

LENGTH: 390

TYPE: PRT

ORGANISM: Homo Sapien

US-09-927-602-4

RESULT 1
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
; CURRENT APPLICATION NUMBER: US/09/927602-4
; PRIORITY FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo Sapien

ALIGMENTS

Scoring table: BL0SUMM2
Gapext 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgri2_6/ptodata/1/pubpaas/US07_PUBCOMB.pep:
2: /cgri2_6/ptodata/1/pubpaas/PCT_NNEW_PUB.pep:
3: /cgri2_6/ptodata/1/pubpaas/US06_NNEW_PUB.pep:
4: /cgri2_6/ptodata/1/pubpaas/US06_PUBCOMB.pep:
5: /cgri2_6/ptodata/1/pubpaas/PCTUS07_NNEW_PUB.pep:
6: /cgri2_6/ptodata/1/pubpaas/US08_NNEW_PUB.pep:
7: /cgri2_6/ptodata/1/pubpaas/US08_PUBCOMB.pep:
8: /cgri2_6/ptodata/1/pubpaas/US08_PUBCOMB.pep:
9: /cgri2_6/ptodata/1/pubpaas/US09_NNEW_PUBCOMB.pep:
10: /cgri2_6/ptodata/1/pubpaas/US09C_PUBCOMB.pep:
11: /cgri2_6/ptodata/1/pubpaas/US09_NNEW_PUB.pep:
12: /cgri2_6/ptodata/1/pubpaas/US10A_PUBCOMB.pep:
13: /cgri2_6/ptodata/1/pubpaas/US10A_PUBCOMB.pep:
14: /cgri2_6/ptodata/1/pubpaas/US10C_PUBCOMB.pep:
15: /cgri2_6/ptodata/1/pubpaas/US10C_PUBCOMB.pep:
16: /cgri2_6/ptodata/1/pubpaas/US10D_PUBCOMB.pep:
17: /cgri2_6/ptodata/1/pubpaas/US10E_PUBCOMB.pep:
18: /cgri2_6/ptodata/1/pubpaas/US10F_PUBCOMB.pep:
19: /cgri2_6/ptodata/1/pubpaas/US11A_PUBCOMB.pep:
20: /cgri2_6/ptodata/1/pubpaas/US11A_NNEW_PUB.pep:
21: /cgri2_6/ptodata/1/pubpaas/US02_NNEW_PUB.pep:
22: /cgri2_6/ptodata/1/pubpaas/US02_PUBCOMB.pep:
RESULT 1
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
; CURRENT INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; INVENTOR: Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927, 602
; CURRENT FILING DATE: 2001-08-09
; PRIORITY FILING NUMBER: US 09/638, 211
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 390

Qy 121 FFWATSRALCSPACSAFFRGTSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 180
 Db 121 FFWATSRALCSPACSAFFRGTSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 180
 Qy 181 NLQVLYPLISDPAINRLVHVRDPRAVLRSREAAAGPILARDNGIVLGTNGKWEADPHL 240
 Db 181 NLQVLYPLISDPAINRLVHVRDPRAVLRSREAAAGPILARDNGIVLGTNGKWEADPHL 240
 Qy 241 RLIREVCRSHVRLAEAATLKKPPLGRGRYLYRVEDLAREPLAEIRALYAFGLTLPQL 300
 Db 241 RLIREVCRSHVRLAEAATLKKPPLGRGRYLYRVEDLAREPLAEIRALYAFGLTLPQL 300
 Qy 301 EAWIINITHSGIGKPIEAFTTSRARNVSQWRHALPFKLLRVOECAQALQIGYR 360
 Db 301 EAWIINITHSGIGKPIEAFTTSRARNVSQWRHALPFKLLRVOECAQALQIGYR 360
 Qy 361 PVIYSDAQDQLTDLVLPRGDHFMSWASPD 390
 Db 361 PVIYSDAQDQLTDLVLPRGDHFMSWASPD 390
 Qy 361 PVIYSDAQDQLTDLVLPRGDHFMSWASPD 390
 Db 361 PVIYSDAQDQLTDLVLPRGDHFMSWASPD 390

RESULT 3
 US-09-927-602-3
 ; Sequence 3 , Application US/09927602
 ; Patent No. US2002006152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Michiko N.
 ; APPLICANT: Akama, Tomoya O.
 ; TITLE OF INVENTION: Methods of Treating Macular Corneal
 ; FILE REFERENCE: P-LJ 4952
 ; CURRENT APPLICATION NUMBER: US/09/927,602
 ; CURRENT FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: US 09/638,211
 ; PRIOR FILING DATE: 2000-08-11
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 3
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic construct
 ; NAME/KEY: VARIANT
 ; LOCATION: (1) .. (395)
 ; OTHER INFORMATION: Xaa = any amino acid

US-09-927-602-3

Query Match 85 7% ; Score 1758 5; DB 9; Length 395;
 Best Local Similarity 87.1%; Pred. No. 6.9e-174;
 Matches 338; Conservative 10; Mismatches 39; Indels 1; Gaps 1;

Qy 1 MWLPRSSKTVVLLAQTCUCLLFIISRGPSPAGEGDHVHVLSSWRGSSFLGQL 60
 Db 1 MWLPRSSKTVVLLAQTCUCLLFIISRGPSPAGEGDHVHVLSSWRGSSFLGQL 59

Qy 61 FSQHPDVFYLMEPAWVWTTLSQGSATLMAVDLDRSFLCDMDVFDAMPQSNSLNA 120
 Db 60 FSQHPDVFYLMEPAWVWTTLSQGSATLMAVDLDRSXFLCDMDVFDAYLPWRRNLSD 119

Qy 121 FENWATSRALCSPACSAFPRTKSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 180
 Db 120 LFQWAVSRALCSPACSAFPRTKSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 179

Qy 120 LFQWAVSRALCSPACSAFPRTKSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 179

Qy 121 FENWATSRALCSPACSAFPRTKSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 180
 Db 121 FENWATSRALCSPACSAFPRTKSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 180
 Qy 121 FENWATSRALCSPACSAFPRTKSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 180
 Db 121 FENWATSRALCSPACSAFPRTKSKQDVCKTLCTROPSLAREACSYSHVVLKEVRRF 180
 Qy 181 NLQVLYPLISDPAINRLVHVRDPRAVLRSREAAAGPILARDNGIVLGTNGKWEADPHL 240
 Db 181 NLQVLYPLISDPAINRLVHVRDPRAVLRSREAAAGPILARDNGIVLGTNGKWEADPHL 240
 Qy 181 NLQVLYPLISDPAINRLVHVRDPRAVLRSREAAAGPILARDNGIVLGTNGKWEADPHL 240
 Db 180 NLQVLYPLISDPAINRLVHVRDPRAVLRSREAAAGPILARDNGIVLGTNGKWEADPHL 239
 Qy 241 RLIREVCRSHVRLAEAATLKKPPLGRGRYLYRVEDLAREPLAEIRALYAFGLTLPQL 300
 Db 240 RLIREVCRSHVRLAEAATLKKPPLGRGRYLYRVEDLAREPLAEIRALYAFGLTLPQL 300
 Qy 301 EAWIINITHSGIGKPIEAFTTSRARNVSQWRHALPFKLLRVOECAQALQIGYR 360
 Db 301 EAWIINITHSGIGKPIEAFTTSRARNVSQWRHALPFKLLRVOECAQALQIGYR 360
 Qy 301 EAWIINITHSGIGKPIEAFTTSRARNVSQWRHALPFKLLRVOECAQALQIGYR 360
 Db 301 EAWIINITHSGIGKPIEAFTTSRARNVSQWRHALPFKLLRVOECAQALQIGYR 360

RESULT 4
 US-09-927-602-2
 ; Sequence 2 , Application US/09927602
 ; Patent No. US2002006152A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Michiko N.
 ; APPLICANT: Akama, Tomoya O.
 ; TITLE OF INVENTION: Methods of Treating Macular Corneal

US-09-927-602-2

Qy 361 PVIYSDAQDQLTDLVLPRGDHFMSWASPD 388
 Db 360 PVIYSDAQDQLTDLVLPRGDHFMSWASPD 387

; TITLE OF INVENTION: Dystrophy
 ; FILE REFERENCE: P-LJ 4852
 ; CURRENT APPLICATION NUMBER: US/09/927,602
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 09/638,211
 ; PRIOR FILING DATE: 2000-08-11
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO: 2
 ; LENGTH: 395
 ; TYPE: PTI
 ; ORGANISM: Homo Sapien
 ; US-09-927-602-2

Query Match 84.3%; Score 1729.5; DB 9; Length 395;
 Best Local Similarity 85.8%; Pred. No. 7.2e-171;
 Matches 33; Conservative 17; Mismatches 37; Indels 1; Gaps 1;

| Query | Subject | Score | DB | Length | Best Local Similarity | Pred. No. | Matches | Conservative | Mismatches | Indels | Gaps | Organism | | | | | | | | |
|-------|---|-------|----|---|-----------------------|-----------|--|--------------|------------|--|------|----------|---|-----|----|---|-----|----|--|-----|
| Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Db | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | | | | | | |
| Qy | 61 FSGQHPDVYLMPEPAWHWTLSQGSSAATLHMAYRDLMSIFLCMDMVDAYMPOSRNLSA | 120 | Db | 60 FNGQHPDVYLMPEPAWHWTLSQGSSAATLHMAYRDLMSIFLCMDMVDAYMPOSRNLSD | 119 | Qy | 121 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 180 | Db | 120 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 179 | Qy | 121 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 180 | Qy | 121 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 180 | | | |
| Qy | 181 NLQVLYPLISDPALNRLTVYLVDPRAVLRSREAAQPTLARDNGIVLGNTGNKVEADPHL | 240 | Db | 180 NLQVLYPLISDPALNRLTVYLVDPRAVLRSREAAQPTLARDNGIVLGNTGNKVEADPHL | 239 | Qy | 241 RLIREVCRSHVRLAEATLKKPPFLGRYKLVRFEDLAREPLABIRALYATFTGLTLPQL | 300 | Db | 240 RVVREVCRSHVRLAEATLKKPPFLGRYKLVRFEDLAREPLABIRALYATFTGLTLPQL | 299 | Qy | 301 EAWHNITHSGIGKPIEAHTTSSRNARNVSQARHALPFKTKLRLQEVCAAGLQLQYR | 360 | Db | 300 EAWHNITHSGIGKPIEAHTTSSRNARNVSQARHALPFKTKLRLQEVCAAGLQLQYR | 359 | Qy | 301 EAWHNITHSGIGKPIEAHTTSSRNARNVSQARHALPFKTKLRLQEVCAAGLQLQYR | 360 |
| Qy | 361 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 388 | Db | 360 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 387 | Qy | 361 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 388 | Db | 360 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 387 | Qy | 361 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 388 | Db | 360 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 387 | | | |

RESULT 5
 US-10-258-080-5
 ; Sequence 5, Application US/10258080
 ; Publication No. US20040029125A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Incyte Genomics, Inc.
 ; APPLICANT: POLICKY, Jennifer L.
 ; APPLICANT: HAFALIA, April J.A.
 ; APPLICANT: BURFORD, Neil
 ; APPLICANT: RING, Huijun Z.
 ; APPLICANT: LAL, Preeta G.
 ; APPLICANT: TRIBOULEY, Catherine M.
 ; APPLICANT: YAO, Monique G.
 ; APPLICANT: YUE, Henry
 ; APPLICANT: ARVIZU, Chandra S.
 ; APPLICANT: DAS, Debopriya
 ; APPLICANT: SANJANAWA, Madhusudan M.
 ; APPLICANT: GANDHI, Ameena R.
 ; APPLICANT: REDDY, Roopa M.
 ; APPLICANT: KHAN, Farrah A.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: RAMKUMAR, Jayalaxmi
 ; APPLICANT: GRIFFIN, Jennifer P.
 ; APPLICANT: AU-YOUNG, Janice K.
 ; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES

; FILE REFERENCE: PI-0070 USN
 ; CURRENT APPLICATION NUMBER: US/10/258,080
 ; CURRENT FILING DATE: 2002-08-15
 ; PRIORITY NUMBER: US 60/203,509
 ; PRIOR FILING DATE: 2000-05-11
 ; PRIORITY NUMBER: US 60/202,234
 ; PRIOR FILING DATE: 2000-05-05
 ; PRIORITY NUMBER: US 60/200,185
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIORITY NUMBER: US 60/198,403
 ; PRIOR FILING DATE: 2000-04-19
 ; PRIORITY NUMBER: PCT/US01/11669
 ; SEQ ID NO: 5
 ; SOFTWARE: PERL Program
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20040029125A1 · 7472777CD1
 ; US-10-258-080-5

Query Match 84.3%; Score 1729.5; DB 15; Length 395;
 Best Local Similarity 85.8%; Pred. No. 7.2e-171;
 Matches 33; Conservative 17; Mismatches 37; Indels 1; Gaps 1;

| Query | Subject | Score | DB | Length | Best Local Similarity | Pred. No. | Matches | Conservative | Mismatches | Indels | Gaps | Organism | | | | | |
|-------|---|-------|----|---|-----------------------|-----------|--|--------------|------------|--|------|----------|---|-----|----|---|-----|
| Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Db | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | Qy | 1 MWLPRFSSKTVVLLAOTCLLFLISRPGPSSPAGEDDRVHVLVLSWRSGSSFVSQL | 60 | | | |
| Qy | 61 FSGQHPDVYLMPEPAWHWTLSQGSSAATLHMAYRDLMSIFLCMDMVDAYMPOSRNLSA | 120 | Db | 60 FNGQHPDVYLMPEPAWHWTLSQGSSAATLHMAYRDLMSIFLCMDMVDAYMPOSRNLSD | 119 | Qy | 121 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 180 | Db | 120 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 179 | Qy | 121 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 180 | Qy | 121 FFWPAWSRALCSPDACSAPRGTSKQDVYCKTLCTROPSLAREACRSYSHVILKEVRF | 180 |
| Qy | 181 NLQVLYPLISDPALNRLTVYLVDPRAVLRSREAAQPTLARDNGIVLGNTGNKVEADPHL | 240 | Db | 180 NLQVLYPLISDPALNRLTVYLVDPRAVLRSREAAQPTLARDNGIVLGNTGNKVEADPHL | 239 | Qy | 241 RLIREVCRSHVRLAEATLKKPPFLGRYKLVRFEDLAREPLABIRALYATFTGLTLPQL | 300 | Db | 240 RVVREVCRSHVRLAEATLKKPPFLGRYKLVRFEDLAREPLABIRALYATFTGLTLPQL | 299 | Qy | 301 EAWHNITHSGIGKPIEAHTTSSRNARNVSQARHALPFKTKLRLQEVCAAGLQLQYR | 360 | Db | 300 EAWHNITHSGIGKPIEAHTTSSRNARNVSQARHALPFKTKLRLQEVCAAGLQLQYR | 359 |
| Qy | 361 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 388 | Db | 360 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 387 | Qy | 361 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 388 | Db | 360 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 387 | Qy | 361 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 388 | Db | 360 PVYSDAQDQRLTDLVLPRGPDHFSWAS | 387 |

RESULT 6
 US-10-648-591-159
 ; Sequence 159, Application US/10648593
 ; Publication No. US20040106132A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
 ; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
 ; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
 ; FILE REFERENCE: D02/3 NP
 ; CURRENT APPLICATION NUMBER: US/10/648,593
 ; CURRENT FILING DATE: 2003-08-26
 ; PRIORITY NUMBER: 60/406,385
 ; PRIOR FILING DATE: 2002-08-27

NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 159
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-159

Query Match 84.3%; Score 1729.5; DB 16; Length 395;
Best Local Similarity 85.8%; Pred. No. 7.2e-171;
Matches 333; Conservative 17; Mismatches 37; Indels 1;

Qy 1 MWLPRFSSKTVTLLAQTCFLIFIRGPSPAGGEDRVVLISSWRSSESSFVQL 60
Db 1 MWLPRVSSATVLLAQ-TELLFLVSRGPSPAGGEARVVLVLESSWRSSESSFVQL 59

Qy 61 FSOHPDVFYLMEPAWHWTTLSQGSAATHMAYRDLNRSI FLCMDVFDAYMPOSRLNSA 120
Db 60 FNQPDVFYLMEPAWHWTTLSQGSAATHMAYRDLNRSI FLCMDVFDAYLPRRNLSD 119

Qy 121 FENWATRALSPPACSPACSFPRGCTSKODVCKTLCTROPSLAREACPSYSHVVKETRFF 180
Db 120 LFQWAVSRALCSPPACSPACSFPRGCTSKODVCKTLCTROPSLAREACPSYSHVVKETRFF 179

Qy 181 NLQVLYPLISDPALNRLVHLVRDPAVLSREAGAPLARDNGIVLTGNGKWEADPHL 240
Db 180 NLQVLYPLISDPALNRLVHLVRDPAVLSREAGAPLARDNGIVLTGNGKWEADPHL 239

Qy 241 RLIREVCRSHVRIAATLKPFPFLGRYRVLRFEDLAREPLAEIRALYAFTGLLTTPQL 300
Db 240 RVYREVCRSHVRIAATLKPFPFLGRYRVLRFEDLAREPLAEIRALYAFTGLSLTPQL 299

Qy 301 EAWHNTHGSGIGKPIFAHTSSRNARNVSOAWRHALPFKYLRYDEVCAAGALQLGYR 360
Db 300 EAWHNTHGSGPGRARBEAKTSRNALNVSOAWRHALPFKYLRYDEVCAAGALQLGYR 359

Qy 361 PVYSDAQORDLTDLVLPQDPDHESWAS 388
Db 360 PVYSEDORNALDLVLPQINGFTWAS 387

RESULT 8
US-10-697-828-7
; Sequence 7, Application US/10697828
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven
; APPLICANT: Lee, Jin Kyu
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, & GST-6
; CURRENT APPLICATION NUMBER: US/10/697,828
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: 60/144,694
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 7
; LENGTH: 395
; TYPE: PRT
; ORGANISM: mus musculus
US-10-697-828-7

Query Match 74.9%; Score 1536.5; DB 16; Length 395;
Best Local Similarity 76.0%; Pred. No. 8.9e-151; Mismatches 62; Indels 1; Gaps 1;

Qy 1 MWLPRFSSKTVTLLAQTCFLIFIRGPSSPAGEDRVVHLVLSWRSGSSFVQL 60
Db 1 MWLPRFSSKTVTLLAQTCFLIFIRGPSSPAGEDRVVHLVLSWRSGSSFVQL 59

Qy 61 FSOHPDVFYLMEPAWHWTTLSQGSAATHMAYRDLNRSI FLCMDVFDAYMPOSRLNSA 120
Db 60 FSQHPDVFYLMEPAWHWTTLSQGSAATHMAYRDLNRSI FLCMDVFDAYLPRRNLSD 119

Qy 121 FENWATRALSPPACSPACSFPRGCTSKODVCKTLCTROPSLAREACPSYSHVVKETRFF 180
Db 120 LFQWAVSRALCSPPACSPACSFPRGCTSKODVCKTLCTROPSLAREACPSYSHVVKETRFF 179

Query Match 84.3%; Score 1729.5; DB 16; Length 395;
Best Local Similarity 85.8%; Pred. No. 7.2e-171;
Matches 333; Conservative 17; Mismatches 37; Indels 1;

Qy 1 MWLPRFSSKTVTLLAQTCFLIFIRGPSSPAGEDRVVHLVLSWRSGSSFVQL 60
Db 1 MWLPRVSSATVLLAQ-TELLFLVSRGPSPAGGEARVVLVLESSWRSSESSFVQL 59

Qy 181 NLQVLYPLISDPALNRLVHLVRDPAVLSREAGAPLARDNGIVLTGNGKWEADPHL 240
Db 180 NLQVLYPLISDPALNRLVHLVRDPAVLSREAGAPLARDNGIVLTGNGKWEADPHL 239

Qy 241 RLIREVCRSHVRIAATLKPFPFLGRYRVLRFEDLAREPLAEIRALYAFTGLLTTPQL 300
Db 240 RVYNEVCRSHVRIAALHKPFPFLQRYRVLRFEDLPLTVRLYAFTRGLGLTPQL 299

Qy 301 EAWINNITHGGIGKPIAEFTSSRNARNVSQAWRHALPFTKLRLRQEVCAGALQLIGYR 360
 Db 300 QTWHNITHGGIPARREAFKTTSRDALSVLPAKIRRVQLCGALQLIGYR 359

Qy 361 PVYSAQQDQRDLTLDLVLPRGDHFSWAS 388
 Db 360 SVHSELEQRDLSDLDLVLPGMDSTKWS 387

RESULT 9

US-10-B11-707-8
 ; Sequence 8, Application US/10841707
 ; Publication No. US20040202649A1

; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Minoru

; TITLE OF INVENTION: Identification of the Meca-79 Antigen

; TITLE OF INVENTION: Methods of Treating L-Selectin-Mediated
 ; Conditions
 ; FILE REFERENCE: P-LJ 4149
 ; CURRENT APPLICATION NUMBER: US/10/841,707
 ; CURRENT FILING DATE: 2004-05-06
 ; PRIOR APPLICATION NUMBER: US/09/569,320A
 ; PRIOR FILING DATE: 2000-05-11
 ; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 8

; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

US-10-B11-707-8

Query Match 74.9%; Score 1516.5; DB 16; Length 395;

Best Local Similarity 76.0%; Pred. No. 8.9e-151; Mismatches 30; Indels 1; Gaps 1;
 Matches 295; Conservative 30; MisMatches -62; Indels 1; Gaps 1;

Qy 1 MWLPRFSSKVNVTULLAQTCLLIFIISRGPSPPAGEDRVHLVLSWRSGSSFLGQI 60
 Db 1 MRLPRFSSSTMISLMLMVQIG-ILVLFLRSQVPSQVPSAGLGERVHLVLSWRSGSSFLGQI 59

Qy 61 FSQHPDVYFLMPEPAHWVWTLSQSGAATLHMAVRDLMRSLFLCDMDVFDAYMPOSRNLSA 120
 Db 60 FSQHPDVYFLMPEPAHWVWTLSQSGAATLHMAVRDLIRSFLCDMDVFDAYLPMRRNISD 119

Qy 121 FENWATSRALCSPPCSAFAFRGTISKQDVCKTLCTROPSFLAREACRSTSYYVHLKEYRFF 180
 Db 120 LFQWAVSRALCSPPVCEAFARNISSEVCPKLCATRGPGLAQACSSYYVHLKEYRFF 179

Qy 181 NLQVLYPLLSDPALNLRIVLVLVRDPRAVLSREAAQPIIARLNGIVLGTNGKVEADPHL 240

Db 180 NLQVLYPLLSDPALNLRIVLVLVRDPRAVLSREAAQPIIARLNGIVLGTNGKVEADPHL 239

Qy 241 RLIRECRSHVRIAATLKKPPFLGRYRLVRFEDLAREPLAETRALYAFGTGLTLPOL 300
 Db 240 RVNNECRSHVRIAATLKKPPFLQDRYLVRFEDLAREPLAETRALYAFGTGLTLPOL 299

Qy 301 EAWINNITHGGIGKPIAEFTSSRNARNVSQAWRHALPFTKLRLRQEVCAGALQLIGYR 360
 Db 300 QTWHNITHGGIPARREAFKTTSRDALSVLPAKIRRVQLCGALQLIGYR 359

Qy 361 PVYSAQQDQRDLTLDLVLPRGDHFSWAS 388
 Db 360 SVHSELEQRDLSDLDLVLPGMDSFKWS 387

RESULT 10
 US-09-927-602-5
 ; Sequence 5, Application US/09927602

; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Michiko N.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 2211

; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-10-408-765A-2211

Query Match 74.3%; Score 1523; DB 9; Length 418;

Best Local Similarity 75.8%; Pred. No. 2.5e-19; Mismatches 62; Indels 2; Gaps 2;

Matches 294; Conservative 30; MisMatches -62; Indels 2; Gaps 2;

Qy 1 MWLPRFSSKTVTULLAQTCLLIFIISRGPSPPAGEDRVHLVLSWRSGSSFLGQI 60
 Db 25 MRLPRFSSSTMISLMLMVQIG-ILVLFLRSQVPSQVPSAGLGERVHLVLSWRSGSSFLGQI 83

Qy 61 FSQHPDVYFLMPEPAHWVWTLSQSGAATLHMAVRDLMRSLFLCDMDVFDAYMPOSRNLSA 120
 Db 84 FSQHPDVYFLMPEPAHWVWTLSQSGAATLHMAVRDLIRSFLCDMDVFDAYLPMRRNISD 143

Qy 121 FENWATSRALCSPPCSAFAFRGTISKQDVCKTLCTROPSFLAREACRSTSYYVHLKEYRFF 180

Db 144 LFQWAVSRALCSPPVCEAFARNISSEVCPKLCATRGPGLADEACSSYYVHLKEYRFF 203

Qy 181 NLQVLYPLLSDPALNLRIVLVLVRDPRAVLSREAAQPIIARLNGIVLGTNGKVEADPHL 240

Db 204 NLQVLYPLLSDPALNLRIVLVLVRDPRAVLSREAAQPIIARLNGIVLGTNGKVEADPHL 263

Qy 241 RLIRECRSHVRIAATLKKPPFLGRYRLVRFEDLAREPLAETRALYAFGTGLTLPOL 300

Db 264 RVNNECRSHVRIAATLKKPPFLQDRYLVRFEDLAREPLAETRALYAFGTGLTLPOL 322

Qy 301 EAHTNITHGGIGKPKBAHTSSRNARNVSQWRHALPFTKLLRQEVCAGALQLIGYR 360

Db 323 QTWHNITHGGIGARRBAPKTTSRDALSVSQAWRHTLPFAKIRRQELCGGALQLIGYR 382

Qy 361 PVYSAQQDQRDLTLDLVLPRGDHFSWAS 388
 Db 383 SVHSELEQRDLSDLDLVLPGMDSFKWS 410

RESULT 11
 US-10-408-765A-2211
 ; Sequence 2211, Application US/10408765A

; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Paby, Eoin D.
 ; APPLICANT: Zhang, Bing Bradford W.
 ; APPLICANT: Gibson, Steven W.

; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 2211

; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-10-408-765A-2211

Query Match 73.8%; Score 1513.5; DB 16; Length 394;
 Best Local Similarity 92.6%; Pred. No. 2.2e-148;
 Matches 289; Conservative 7; Mismatches 15; Indels 1; Gaps 1;

1 MWLPPRFSKTVTULLIAQTCVTLFISRPGPSSPAGGEPRVHVLSSWRSGSSFLGQI 60
 1 MWLPPRFSKTVTULLIAQ-TULLFISRPGPSSPAGGEPRVHVLSSWRSGSSFLGQI 59
 61 FSQHPDVFYLMEPANHWTLSQGSATLIMAVDLMRSLFICMDVDFAYMPQSRLNSA 120
 60 FNQHPDVFYLMEPANHWTLSQGSATLIMAVDLVRSLFICMDVDFAYLPWNRNLSD 119

RESULT 13
 US-10-007-262-1
 ; Sequence 1 Application US/10007262
 ; Publication No. US20020164748A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bistrup, Annette
 ; APPLICANT: Roser, Steven D.
 ; APPLICANT: Tangemann, Kiraten
 ; APPLICANT: Hemmerich, Stefan
 ; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
 ; FILE REFERENCE: 651-0-107CIP
 ; CURRENT APPLICATION NUMBER: US/10/007,262
 ; CURRENT FILING DATE: 2001-11-08
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 1
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: H. sapiens
 US-10-007-262-1

Query Match 49.1%; Score 1008; DB 13; Length 386;
 Best Local Similarity 52.3%; Pred. No. 9.1e-36;
 Matches 56; Mismatches 104; Indels 28; Gaps 7;

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 14 LLLAQTCVTLIFLIFSP-----GPSSPAGEDDRVHVLSSWRSGSSFLGQ 59
 1 MLLPKCKMFLFLVSQMAILALFFHMYSNIISSMSKAOPERMHVVLSSWRSGSFRVGQ 60

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 60 LFGQHDDVFYLMEPANHWTLSQGAATLIMAVDLMRSLFICMDVDFAYM-POSRLN 118
 61 LFGQHDDVFYLMEPANHWTLSQGAATLIMAVDLMRSLFICMDVDFAYM-POSRLN 120

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 119 SAFFENATSRALCSPPACSAFPGTISKODVCKTLCTROPSLARACRSYSHVVLKEVR 178
 121 SSLFOENSRALCSACACDIIPODEIIPRAHCRLLCSSQPFPEVVEACRSYSHVVLKEVR 180

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 121 SSLFOENSRALCSACACDIIPODEIIPRAHCRLLCSSQPFPEVVEACRSYSHVVLKEVR 180

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 179 FPNLQLQYPLISDPAINLRIVHLVRDPAVRSREAGPILARDNGVILGN-GKVEAD 237

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 181 FFNLQSLYPLIKDPSNLHIVLVRDPAVRSRERTKGDLMDSRVMGHEQKLKED 240

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 14 LLLAQTCVTLIFLIFSP-----GPSSPAGEDDRVHVLSSWRSGSSFLGQ 59
 1 MLLPKCKMFLFLVSQMAILALFFHMYSNIISSMSKAOPERMHVVLSSWRSGSFRVGQ 60

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 60 LFGQHDDVFYLMEPANHWTLSQGAATLIMAVDLMRSLFICMDVDFAYM-POSRLN 118
 61 LFGQHDDVFYLMEPANHWTLSQGAATLIMAVDLMRSLFICMDVDFAYM-POSRLN 120

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 119 SAFFENATSRALCSPPACSAFPGTISKODVCKTLCTROPSLARACRSYSHVVLKEVR 178
 121 SSLFOENSRALCSACACDIIPODEIIPRAHCRLLCSSQPFPEVVEACRSYSHVVLKEVR 180

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 179 FPNLQLQYPLISDPAINLRIVHLVRDPAVRSREAGPILARDNGVILGN-GKVEAD 237

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 181 FFNLQSLYPLIKDPSNLHIVLVRDPAVRSRERTKGDLMDSRVMGHEQKLKED 240

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 238 PHRLRREVCRSHVRAEAATLKP-PFPLRGYRYLRFEDIAREPLAEIRALYAFGTGLT 296

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 241 QPYYNNQVICOSLEYK--TQSLPKALQERYLRYLRFEDIAREPLAEIRALYAFGTGLT 298

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 297 TQPLEWIHNTHGSGIGKPKLEAFTTSRNARNVSQAWRHALPFTKLRVQEVCAQALQ 356

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 299 LPFLQHDTWVNNTTRGKNGD--HAFTNARDALNVSNQAWRWSLYPEKVSRLQKACGDAMNL 356

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 357 LGYRPVYSAFQDQRLTDLVLPRGPDFSWASPD 390

Qy | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 357 LGYRHYRSEQBORNLLIDLL----STWTYPE 383

RESULT 14
 US-10-841-707-6
 ; Sequence 6 Application US/10841707
 ; Publication No. US200402669A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Minoru
 ; APPLICANT: Fukuda, Minoru

179 FPNLQLQYPLISDPAINLRIVHLVRDPAVRSREAGPILARDNGVILGN-GKVEAD 237
 181 FFNLQSLYPLIKDPSNLHIVLVRDPAVRSRERTKGDLMDSRVMGHEQKLKED 240

238 PHRLRREVCRSHVRAEAATLKP-PFPLRGYRYLRFEDIAREPLAEIRALYAFGTGLT 296

APPLICANT: Yeh, Jiunn-Chern
 APPLICANT: Hirao, Nobuyoshi
 TITLE OF INVENTION: Identification of the Meca-79 Antigen
 TITLE OF INVENTION: and Related Methods of Treating L-Selectin-Mediated
 Conditions
 FILE REFERENCE: P-LJ 4149
 CURRENT APPLICATION NUMBER: US/10/841,707
 CURRENT FILING DATE: 2004-05-06
 PRIOR APPLICATION NUMBER: US/05/569,320A
 PRIOR FILING DATE: 2000-05-11
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 6
 LENGTH: 380
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-841-707-6

Query Match 48.9%; Score 1003.5; DB 16; Length 380;
 Best Local Similarity 53.1%; Pred. No. 2.6e-95;
 Matches 206; Conservative 55; Mismatches 102; Indels 25; Gaps 8;

Qy 13 VLLAQTCILF-----IISRGPGSSPAGGEDDRVYLVLSWRSQGSFLGOLFQHP 65
 S LFLVSOMAILALFFMYSHNISLSSMKAQD----ERMHVLVLSWRSQGSFVGQFLGGHP 60

Db 66 DVFYLMEPAWHWTLSQGSATLHMAVRDLMRSLFLCDMDVFDAYM-POSRLSAFFW 124
 Db 61 DVFYLMEPAWHWTLPKQSTAWLHMAVRDLIRAVFLCDMSVFDAYMEPPROSSLFW 120

Qy 125 ATSRALCSPACSAFPRTISKODVCKTLCTROPFLSAREACRSYSHVYLKEVRFNLYQV 184
 Db 121 ENSRAICSAACDIPQDEBTPRANCILLSQQPPEVVEACRSYSHVYLKEVRFNLYQ 180

Qy 185 LYPLSDPANLRIVHLVRPRAVLRSREAGPILARDNGCIVLGTN-GKWEADPHRLI 243
 Db 181 LYPLKDPSNLNHLVHLVRPRAVRSRETKGDLMDSIIVMGHEQKIKKEQPYTYM 240

Qy 244 RECRVSHVRAEAATLKP-PPFLRGTRLYRFEDLAREPLAEIRALYAFFGLTLPQLEA 302
 Db 241 QVICQSLEYK-TIQSLPKALQBRYLRYEDLARAPTAQTSRMYEFFGLEPLPHOT 298

Qy 303 WIHNTHGSGIGKPIEAFHTSSRNARNVSDAHRALPFTKILRVQBCAGALQGYREV 362
 Db 299 WHNNTRGKGMD-HAFHTNDAVNLNQVSLPVEVSRSLQKACGAMNLGYREV 356

Qy 363 YSDQDQRDLTDLVLRGPDPHFSWASPD 390
 Db 357 RSEQBORNLLDLL-----STWTYPE 377

Search completed: June 23, 2005, 09:27:53
 Job time : 50.3029 secs

RESULT 15

US-10-427-631-11

Sequence 11, Application US/10427631
 Publication No. US20030175923A1

GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
 CORLEY, Neil C.; GUEGLER, Karl J.;
 BAUGH, Mariah R.; LAL, Preeti G.;
 YUE, Henry; HILLMAN, Jennifer L.;
 AZIMZAI, Valda

TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS

FILE REFERENCE: PP-059-1 DIV
 CURRENT APPLICATION NUMBER: US/10/427,631
 CURRENT FILING DATE: 2003-04-29
 PRIOR APPLICATION NUMBER: US 09/786,240
 PRIOR FILING DATE: 2002-03-12
 PRIOR APPLICATION NUMBER: PCT/US99/20989
 PRIOR FILING DATE: 1999-09-09
 PRIOR APPLICATION NUMBER: US 60/172,220
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: US 60/155,248
 PRIOR FILING DATE: 1998-11-04

PRIOR APPLICATION NUMBER: US 60/133,642
 PRIOR FILING DATE: 1999-05-11
 NUMBER OF SEQ ID NOS: 33
 SOFTWARE: PERL Program
 SEQ ID NO: 11
 LENGTH: 386
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. US20030175923A1 2617407CD1
 US-10-427-631-11

| Query | Match | Length | Score | Best Local Similarity | Pred. No. | Gaps | |
|-------|--|--------|--------------------------------|--------------------------------|-----------|----------|--|
| Qy | 14 LLLAQTCILFISRP-----GPSSPAGEDDRVYLVLSWRSQSSPFLQ 59 | 386 | 970; | DB 14; | 51.3%; | 7; | |
| | 1 MLLPKKMLFLVSQMAILALFFMHSHNISLSSMKAQPERMHLVLSWRSQSSPFLQ 60 | 386 | GPSSPAGEDDRVYLVLSWRSQSSPFLQ 59 | 51.3%; | 8.3e-92; | | |
| Db | 60 LFSQHPDVYLMEPAWHWTLSQGSATLHMAVRDLRSLFLCDMDVFDAYM-PQSRLN 118 | 386 | 59 | DB 14; | 51.3%; | 7; | |
| | 61 LGQHPDVYLMEPAWHWTMFQSTAWLHMAVRDLIRAVFLCDMSVFDAYMEPPRQ 120 | 386 | GPSSPAGEDDRVYLVLSWRSQSSPFLQ 59 | 51.3%; | 8.3e-92; | | |
| Qy | 69 SAFFNWATSRALCSPACSAFPRTISKODVCKTLCTROPFLSAREACRSYSHVYLKEVR 178 | 386 | 59 | DB 14; | 51.3%; | 7; | |
| | 121 SSUHQWENSRALCSAPACDIPQDESSPQLTAGSCAVNSPLKLUKACRSYSHVYLKEVR 180 | 386 | 59 | GPSSPAGEDDRVYLVLSWRSQSSPFLQ 59 | 51.3%; | 8.3e-92; | |
| Db | 129 FPNLQVLYLPSDPAUNRIVHLVRDPRAVLRSREAGPILARDNGCIVLGTN-GKWEAD 237 | 386 | 59 | DB 14; | 51.3%; | 7; | |
| | 181 FPNLQSLPILKOPSLNLHVHLVRDPRAVFSSRTKGDLMDSRIVNGQEQLKKED 240 | 386 | 59 | GPSSPAGEDDRVYLVLSWRSQSSPFLQ 59 | 51.3%; | 8.3e-92; | |
| Qy | 238 PHPLFLIREYCRSHVRAEAATLKP-PPFLRGTRLYRFEDLAREPLAEIRALYAFFGLTLPQL 296 | 386 | 59 | DB 14; | 51.3%; | 7; | |
| | 241 QPYVVMQVICQSQIYK-TIQSLPKALQERTLIVRVEDLARAPVQTTSRMYEFVGLEF 298 | 386 | 59 | GPSSPAGEDDRVYLVLSWRSQSSPFLQ 59 | 51.3%; | 8.3e-92; | |
| Db | 297 TPQLEAWIHNITHGSGIGKPIEAFHTSSRNARNVSDAHRALPFTKILRVQEVCAAGALQL 356 | 386 | 59 | DB 14; | 51.3%; | 7; | |
| | 299 LPHIQTWVNITROKGMD-HAFHTNDAVNLNQVSLPVEVSRSLQKACGDMNL 356 | 386 | 59 | GPSSPAGEDDRVYLVLSWRSQSSPFLQ 59 | 51.3%; | 8.3e-92; | |

THIS PAGE BLANK (USP TO)

RESULT 2
US-0-949-016-8893
; Sequence 8893, Application US/0949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 7011
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7011

Query Match 100.0%; Score 2051; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.9e-217; Indels 0;
Matches 390; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MWLPRFSSKTVVLLAQTCILFLISRPGPSSPAGEDRHYVLSSMRSGSSFLGQL 60
Db 42 MWLPRFSSKTVVLLAQTCILFLISRPGPSSPAGEDRHYVLSSMRSGSSFLGQL 101
Qy 61 FSQHDPDYVLMEPAWHWTLSQGSAATLHMAYRDLMRSFLCDMDVDAYMPQRNLSA 120
Db 102 FSQHDPDYVLMEPAWHWTLSQGSAATLHMAYRDLMRSFLCDMDVDAYMPQRNLSA 161
Qy 121 FENWATSRALCSPACSAFRGTISKODVCTLCTRQPFSLAREACRSYSHVVLKEVRFF 180
Db 162 FENWATSRALCSPACSAFRGTISKODVCTLCTRQPFSLAREACRSYSHVVLKEVRFF 221
Qy 181 NLQVLYPLISDPAHLNRYVHLVDRPVLRSREAGPILARDNGIVLGTNGKWEADPHL 240
Db 222 NLQVLYPLISDPAHLNRYVHLVDRPVLRSREAGPILARDNGIVLGTNGKWEADPHL 281
Qy 241 FLIREYCRSHVRIAATLKPPFLGRYRVLRFEDLAREPLAEIRALYAFGLTLTPOL 300
Db 282 FLIREYCRSHVRIAATLKPPFLGRYRVLRFEDLAREPLAEIRALYAFGLTLTPOL 341
Qy 301 FAWIHNTHGGIGKPEAFTTSSRNARNYQAWRHLPFTKILRVOECA GALQOLGYR 360
Db 342 FAWIHNTHGGIGKPEAFTTSSRNARNYQAWRHLPFTKILRVOECA GALQOLGYR 401
Qy 361 PYVSADQQRDTIDLVLRPGDHFNSASPD 390
Db 402 PYVSADQQRDTIDLVLRPGDHFNSASPD 431

RESULT 3
US-09-949-016-7011
; Sequence 7011, Application US/0949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-045-284A-2

Query Match 84.3%; Score 1729.5; DB 4; Length 395;
Best Local Similarity 85.8%; Pred. No. 7.4e-182; Indels 1; Gaps 1;
Matches 333; Conservative 17; Mismatches 37;

Qy 1 MWLPRFSSKTVVLLAQTCILFLISRPGPSSPAGEDRHYVLSSMRSGSSFLGQL 60
Db 1 MWLPRFSSKTVVLLAQ-TFLLIPLVSRGPSSPAGEDRHYVLSSMRSGSSFLGQL 59
Qy 61 FSQHDPDYVLMEPAWHWTLSQGSAATLHMAYRDLMRSFLCDMDVDAYMPQRNLSA 120
Db 60 FNOHQHDPDYVLMEPAWHWTLSQGSAATLHMAYRDLMRSFLCDMDVDAYMPQRNLSD 119
Qy 121 FENWATSRALCSPACSAFRGTISKODVCTLCTRQPFSLAREACRSYSHVVLKEVRFF 180
Db 120 LFQWAVSRALCSPACSAFRGTISKODVCTLCTRQPFSLAREACRSYSHVVLKEVRFF 179
Qy 181 NLQVLYPLISDPAHLNRYVHLVDRPVLRSREAGPILARDNGIVLGTNGKWEADPHL 240
Db 180 NLQVLYPLISDPAHLNRYVHLVDRPVLRSREAGPILARDNGIVLGTNGKWEADPHL 239
Qy 241 RLIREYCRSHVRIAATLKPPFLGRYRVLRFEDLAREPLAEIRALYAFGLTLTPOL 300
Db 240 RVVREYCRSHVRIAATLKPPFLGRYRVLRFEDLAREPLAEIRALYAFGLSLTPOL 299
Qy 301 EAWIHTHGGIGKPIEAFTTSSRNARNYQAWRHLPFTKILRVOECA GALQOLGYR 360
Db 300 EAWIHTHGGIGKPIARAKTSSRNALNYQAWRHLPFAKIRRVQELCAGALQOLGYR 359
Qy 361 PVYSAQDQRLTDLVLPRGPDPHFSWA S 388
Db 360 PVYSEDEQRNLALDLVLPRGLNGFTWA S 387

RESULT 4
US-09-045-284A-2
; Sequence 2, Application US/0945284A
; Patent No. 6265192
; GENERAL INFORMATION:
; APPLICANT: Bibirup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCIN SULFOTRANSFERASE -3
; FILE REFERENCE: 6510-1070US1
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-045-284A-2

Query Match 49.1%; Score 1008; DB 2; Length 386;
Best Local Similarity 52.3%; Pred. No. 2.7e-102;

Matches 206; Conservative 56; Mismatches 104; Indels 28; Gaps 7; Qy 14 LLLAQTTCUFLFISRP-----GPSSPAGGEDRVTYVLUSSWRSGSSSPFGQ 59 Db 1 MLLPKMKMFLFLPSQMAILALPFPHMYSHNISLMSKAOPERMVLVLUSSWRSGSSSPFGQ 60 Qy 60 LFSQHDPVYLMEPAWHVNITLSSQGSAATHMAYRDLMRSIFLCMDVFDAYM-PQSRLN 118 Db 61 LFGQHPDVYLMEPAWHVNITLSSQGSAATHMAYRDLMRSIFLCMDVFDAYM-PGPRQ 120 Qy 119 SAffENPATSLCSPPACSAFPRTISKODVCKTLCTROPSLAREACSYSHVVLKEYR 178 Db 121 SLLFOWENSLCSAPACDIPQDEIIPRAHCRLCSQQPEVVEKACRSYSHVVLKEYR 180 Qy 179 FENLQVLYPLPSDPAHLNRLVYLRDPRAVLRSREAAQPLARDNGIVLGTN-GKWEAD 237 Db 181 FENLQVLYPLPSDPAHLNRLVYLRDPRAVLRSREAAQPLARDNGIVLGTN-GKWEAD 240 Qy 238 PHRLIREVCRSHVRIABAATLKP-PPLFRGRYRLVRPEDLAREPLAEIRALYAFTGLTL 296 Db 241 QPYVMQVICQSQLEIYK-TIQSLPKALQERYLVRYEDLARAPVQTSRMTEFVGLEF 298 Qy 297 TPOLEAWHINITHSGIGRPIEAFTTSSNARYNSQAMRHALPTKILRVQEVCAAGALQ 356 Db 299 LPHLQTVHNITRGKGMGD-HAFHTNARDALNVSQAMRWSLPYEVKSRQLQACGDAQNL 356 Qy 357 LGYRPPVYSAQDQDILTDLYLPRGPDHFWSWASPD 390 Db 357 LGYRHRVRSBEBQRNLDDL-----STWTVPE 383

RESULT 6
 US-09-786-240-11
 ; Sequence 11; Application US/09786240
 ; Patent No. 6558335
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE PHARMACEUTICALS, INC.
 ; TANG, Y. Tom
 ; CORLEY, Neil C.
 ; BAUGHN, Karl J.
 ; BAUGHN, Mariah R.
 ; LAL, Preeti
 ; YUE, Henry
 ; HILLMAN, Jennifer L.
 ; APPLICANT: AZIMZAI, Yalda
 ; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 ; FILE REFERENCE: PF-0592 PCT
 ; CURRENT APPLICATION NUMBER: US/09/786 , 240
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
 ; NUMBER OF SEQ ID NOS: 33
 ; SEQ ID NO 11
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. 6558935 2617407CD1
 ; US-09-786-240-11

Query Match 47.3%; Score 970; DB 4; Length 386;
 Best Local Similarity 51.3%; Pred. No. 4.3e-98;
 Matches 202; Conservative 52; Mismatches 112; Indels 28; Gaps 7;

Qy 14 LLLAQTTCUFLFISRP-----
 Db 1 MLLPKMKMFLFLPSQMAILALPFPHMYSHNISLMSKAOPERMVLVLUSSWRSGSSSPFGQ 60
 Qy 60 LFSQHDPVYLMEPAWHVNITLSSQGSAATHMAYRDLMRSIFLCMDVFDAYM-PQSRLN 118
 Db 61 LFGQHPDVYLMEPAWHVNITLSSQGSAATHMAYRDLMRSIFLCMDVFDAYM-PGPRQ 120
 Qy 119 SAffENPATSLCSPPACSAFPRTISKODVCKTLCTROPSLAREACSYSHVVLKEYR 178
 Db 121 SLLFOWENSLCSAPACDIPQDEIIPRAHCRLCSQQPEVVEKACRSYSHVVLKEYR 180
 Qy 179 FENLQVLYPLPSDPAHLNRLVYLRDPRAVLRSREAAQPLARDNGIVLGTN-GKWEAD 237
 Db 181 FENLQVLYPLPSDPAHLNRLVYLRDPRAVLRSREAAQPLARDNGIVLGTN-GKWEAD 240
 Qy 238 PHRLIREVCRSHVRIABAATLKP-PPLFRGRYRLVRPEDLAREPLAEIRALYAFTGLTL 296
 Db 241 QPYVMQVICQSQLEIYK-TIQSLPKALQERYLVRYEDLARAPVQTSRMTEFVGLEF 298
 Qy 297 TPOLEAWHINITHSGIGRPIEAFTTSSNARYNSQAMRHALPTKILRVQEVCAAGALQ 356
 Db 299 LPHLQTVHNITRGKGMGD-HAFHTNARDALNVSQAMRWSLPYEVKSRQLQACGDAQNL 356

RESULT 5
 US-09-190-911-1
 ; Sequence 1; Application US/09190911
 ; Patent No. 6365365
 ; GENERAL INFORMATION:
 ; APPLICANT: Bistrup, Annette
 ; APPLICANT: Tangemann, Kirten
 ; APPLICANT: Hemmerich, Stefan
 ; APPLICANT: Rosen, Steven D.
 ; FILE REFERENCE: 6510-107CIP
 ; CURRENT APPLICATION NUMBER: US/09/190,911
 ; CURRENT FILING DATE: 1998-11-12
 ; EARLIER APPLICATION NUMBER: 09/045,284
 ; EARLIER FILING DATE: 1998-03-20
 ; SOFTWARE: fastSEQ for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: H. sapiens
 ; US-09-190-911-1

Query Match 49.1%; Score 1008; DB 3; Length 386;
 Best Local Similarity 52.3%; Pred. No. 2.7e-102;
 Matches 206; Conservative 56; Mismatches 104; Indels 28; Gaps 7;

Qy 14 LLLAQTTCUFLFISRP-----
 Db 1 MLLPKMKMFLFLPSQMAILALPFPHMYSHNISLMSKAOPERMVLVLUSSWRSGSSSPFGQ 60
 Qy 60 LFSQHDPVYLMEPAWHVNITLSSQGSAATHMAYRDLMRSIFLCMDVFDAYM-PQSRLN 118
 Db 61 LFGQHPDVYLMEPAWHVNITLSSQGSAATHMAYRDLMRSIFLCMDVFDAYM-PGPRQ 120
 Qy 119 SAffENPATSLCSPPACSAFPRTISKODVCKTLCTROPSLAREACSYSHVVLKEYR 178
 Db 121 SLLFOWENSLCSAPACDIPQDEIIPRAHCRLCSQQPEVVEKACRSYSHVVLKEYR 180
 Qy 179 FENLQVLYPLPSDPAHLNRLVYLRDPRAVLRSREAAQPLARDNGIVLGTN-GKWEAD 237
 Db 181 FENLQVLYPLPSDPAHLNRLVYLRDPRAVLRSREAAQPLARDNGIVLGTN-GKWEAD 240

RESULT 7

Qy 357 LGYRPPVYSAQQDIDLTDLVLPRGPDPDFSWASPD 390
 Db 357 LGYRHVRSEFQRNLLDIL-----STWIVPE 383

GENERAL INFORMATION:
 APPLICANT: Uchimura, Kenji
 APPLICANT: Muramatsu, Hideki
 APPLICANT: Kadomatsu, Kenji
 APPLICANT: Kannagi, Reiji
 APPLICANT: Habuchi, Osami
 APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACTYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND FILE REFERENCE: TOYAMA1.00IAUS

CURRENT APPLICATION NUMBER: US/09/471, 867
 CURRENT FILING DATE: 1999-12-23
 EARLIER APPLICATION NUMBER: US/09/263, 023
 EARLIER FILING DATE: 1999-03-05
 EARLIER APPLICATION NUMBER: JP 10-54007
 EARLIER FILING DATE: 1998-03-05
 EARLIER APPLICATION NUMBER: JP 10-177844
 EARLIER FILING DATE: 1998-06-24
 EARLIER APPLICATION NUMBER: JP 10-177844
 EARLIER FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 10
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO: 2
 LENGTH: 483
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-471-867-2

Query Match 32.4%; Score 665; DB 4; Length 483;
 Best Local Similarity 40.4%; Pred. No. 2.e-64;
 Matches 150; Conservative 64; Mismatches 121; Indels 36; Gaps 11;

Qy 37 GGEDRPHVILVSSWRGGSSPFGQLESQHDVFYLMEPAWHVVTTLSQGSATLHMARDL 96
 Db 113 GGDKRQLQVYVPTWRRGSSPFGEFLFNQNPEWFLYEPVNWHWQKLYPGDAVSLOGARDM 172

Qy 97 MRSIFLCMDYFDAYMPQ--SRNLS--AFFNWATSPACSPACSFPRSTISKD--V 149
 Db 173 LSALYRCDSLVSFQLSPAGSGRNLTLGFGAATKVKVCSPLCPAYKEVVGLYDRV 232

Query Match 32.4%; Score 665; DB 3; Length 483;
 Best Local Similarity 40.4%; Pred. No. 2.e-64;
 Matches 150; Conservative 64; Mismatches 121; Indels 36; Gaps 11;

Qy 37 GGEDRPHVILVSSWRGGSSPFGQLESQHDVFYLMEPAWHVVTTLSQGSATLHMARDL 96
 Db 113 GGDKRQLQVYVPTWRRGSSPFGEFLFNQNPEWFLYEPVNWHWQKLYPGDAVSLOGARDM 172

Qy 97 MRSIFLCMDYFDAYMPQ--SRNLS--AFFNWATSPACSPACSFPRGCTISKD--V 149
 Db 173 LSALYRCDSLVSFQLSPAGSGRNLTLGFGAATKVKVCSPLCPAYKEVVGLYDRV 232

Qy 210 RS-----REAGAPILLARD-----NGIVLGT--NGKMYEADPH-LRLIREV 246
 Db 292 SSRIRSRHGLIRESLQVRSRSDPRAHMPFLEAAHGKLGAKCEGMCGPADDYHALGMEVI 351

Qy 247 CRSHVRIAEATLKPPFLGRYRLVRFIDAREPLAEIRALYAFTGTTTPQLEWIHN 306
 Db 352 CNSMAKTLQTA-LQPPDWLQGHYLVTVRYEDLVGDPFKTLRVDFGLVSPMEMOFALN 410

Qy 150 CKTLCTROPSLAREACRSYSHVILKEYRFPNLQVLYLSDPALNRIVHLVYDPAVL 209
 Db 233 CKK-CPPQLRFEERCKRTVVKIGTRVFDAYVLPJLKDVKVHLVRDPAVA 291

Qy 210 RS-----REAGAPILLARD-----NGIVLGT--NGKMYEADPH-LRLIREV 246
 Db 292 SSRIRSRHGLIRESLQVRSRSDPRAHMPFLEAAHGKLGAKCEGMCGPADDYHALGMEVI 351

Qy 247 CRSHVRIAEATLKPPFLGRYRLVRFIDAREPLAEIRALYAFTGTTTPQLEWIHN 306
 Db 352 CNSMAKTLQTA-LQPPDWLQGHYLVTVRYEDLVGDPFKTLRVDFGLVSPMEMOFALN 410

RESULT 9

Qy 307 ITHGSG-IGKPIEAFTHSRARNVNSQMRHALPFTKILRQEVCAGALQLGYRPVYSA 365
 Db 411 MTSGSGSSSKP---FVVSARNATQAANAWRTALFQOIKQVEEFCYQPMAVLGYERVNSP 467

Qy 366 DQQRDLTLDLV 376
 Db 468 BEVKDLSTKL 478

RESULT 8

Qy 307 ITHGSG-IGKPIEAFTHSRARNVNSQMRHALPFTKILRQEVCAGALQLGYRPVYSA 365
 Db 411 MTSGSGSSSKP---FVVSARNATQAANAWRTALFQOIKQVEEFCYQPMAVLGYERVNSP 467

Qy 366 DQQRDLTLDLV 376
 Db 468 BEVKDLSTKL 478

GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: C1001307
 CURRENT APPLICATION NUMBER: US/09/949, 016
 PRIORITY NUMBER: 6812339
 PRIORITY FILING DATE: 2000-04-14
 PRIORITY APPLICATION NUMBER: 60/241, 755
 PRIORITY FILING DATE: 2000-10-02
 PRIORITY APPLICATION NUMBER: 60/237, 768
 PRIORITY FILING DATE: 2000-10-03
 PRIORITY APPLICATION NUMBER: 60/231, 498
 PRIORITY FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: For Windows Version 4.0
; SEQ ID NO: 9449
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9449

Query Match Score 659.5; DB 4; Length 608;
Best Local Similarity 39.1%; Pred. No. 1.5e-63;
Matches 151; Conservative 68; Mismatches 122; Indels 45; Gaps 12;

Qy 31 GPSSPAG-----GEDRVRHVLSSWRSGSSFLGOLFSDQHDFVYLMEPAWHVVTTL 81
Db 223 GVAAPGGNGTRGTGGVGDKRQLQYVFTWRSGSSPFGEFLQVFFPEVYPMWVQKL 282

Qy 82 SGQSAATLIMAARDLMSFLCDMDVFDAYMPQ--SRNLS-AFFNATSRALCSPAC 136
Db 283 YPGDASVLSQGAARDLMSALYRCDSVFQYSPAGSGGRNLTTLGIGAATNKVVCSSPLC 342

Qy 137 SAFAFRGTISKQD--VCKTLCTRQPSLAREACRSYSHVVLKEVRFNNIQLVYPLLSDPAL 194
Db 343 PAYKEVVGDDRVCKK-CPPQBLARFEECYRTLVIGYTRVEDAVLPLRDALLLRDBAL 401

Qy 195 NLRIYHVLYDRPRAVLRS-----REAGAPTLARD-----NGIVLTGNGKWW 234
Db 402 DLKVTHVDRPRAVASSRIRSRHLIRESLQVVRSDPRAHNPFLEA GHKGAKCGV 461

Qy 235 --EADPH-LRLIREVCRSHVRIAATLKPPFLGRYRLVRFEDLAREPLAIRALYAF 291
Db 462 GGPADYHALGAMEVICNSMAKTLQTA-LQPDNMQGHLYLVRVRLDLYGDPVKTLRRVYDF 520

Qy 292 TGLTLTPOLEAWHTNITRGSG-IGKPIEAFTHSERNARNVSQAWRHLPFKILRVQEVIC 350
Db 521 VGLLUSPENQFLAUNNTSSGSSSKP--FVVSARNATOAAANAWRALTQFOQIKQVBBFC 577

Qy 351 AGAQLQLGYRPVYSSADQQRDLTDLV 376
Db 578 YQPAVGLGYERVNPEEVKDSLTKL 603

RESULT 11
US-09-471-867-4
; Sequence 4, Application US/09471867
; Patent No. 6455389

GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Kadomatsu, Hideki
; APPLICANT: Kamagai, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAM41_001AU5
; CURRENT APPLICATION NUMBER: US 09/471,867
; CURRENT FILING DATE: 1999-12-23
; EARLIER APPLICATION NUMBER: US 09/263,023
; EARLIER FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: JP 10-54007
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: JP 10-177844
; EARLIER FILING DATE: 1998-06-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-471-867-4

Query Match Score 656.5; DB 4; Length 484;
Best Local Similarity 39.7%; Pred. No. 2.3e-63;
Matches 149; Conservative 68; Mismatches 121; Indels 37; Gaps 12;

Qy 34 SPAGGEDEPRH-VLVLSSWRSGSSFLGOLFSDQHDFVYLMEPAWHVVTTL 92
Db 110 APEVGDRGRHMMVFTWRSGSSPFGELENQNPEVFLYEPVHVKLYPFDVSLQGA 169

Qy 93 VRDLMRSIFLCMDVFDAYMPQ--SRNLS-AFFNATSRALCSPACSAFRGTSKQ 147
Db 170 ARDMLSLYRCDSVFQYSPAGSGGRNLTTLGIGAATNKVVCSSPLCPAYKEVVL 229

Query Match Score 656.5; DB 3; Length 484;
Best Local Similarity 39.7%; Pred. No. 2.3e-63;
Matches 149; Conservative 68; Mismatches 121; Indels 37; Gaps 12;

Qy 148 D--VCKTLCTRQPSLAREACRSYSHVVLKEVRFNNIQLVYPLLSDPALNRLVLRD 205
Db 230 DDRVCKK-CPPQBLARFEECYRTLVIGYTRVEDAVLPLRDALLLRDBAL 288

RESULT 10
US-09-263-023-4
; Sequence 4, Application US/09263023
; Patent No. 6037159

GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Kamagai, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAM41_001AU5
; CURRENT APPLICATION NUMBER: US 09/263,023
; CURRENT FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: JP 10-54007
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: JP 10-177844
; EARLIER FILING DATE: 1998-06-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-263-023-4

Query Match Score 656.5; DB 4; Length 484;
Best Local Similarity 39.7%; Pred. No. 2.3e-63;
Matches 149; Conservative 68; Mismatches 121; Indels 37; Gaps 12;

Qy 34 SPAGGEDEPRH-VLVLSSWRSGSSFLGOLFSDQHDFVYLMEPAWHVVTTL 92
Db 110 APEVGDRGRHMMVFTWRSGSSPFGELENQNPEVFLYEPVHVKLYPFDVSLQGA 169

Qy 93 VRDLMRSIFLCMDVFDAYMPQ--SRNLS-AFFNATSRALCSPACSAFRGTSKQ 147
Db 170 ARDMLSLYRCDSVFQYSPAGSGGRNLTTLGIGAATNKVVCSSPLCPAYKEVVL 229

Query Match Score 656.5; DB 3; Length 484;
Best Local Similarity 39.7%; Pred. No. 2.3e-63;
Matches 149; Conservative 68; Mismatches 121; Indels 37; Gaps 12;

Db 230 DDRVCKK-CPPQLARFEECRKYRTLVIKVGRVFDVAVLPLRDPALDKVTHLYRDP 288
 Qy 206 RAVLRS-----REAGPILLARD-----NGILVLTGNGKRV-EADPH-LRL 242
 Db 289 RAVASSRTRSRHGLIRESQVRDRPRAHMPFLEAAGHKLGAKEVGPGDHYAIGA 348
 Result 13
 US-08-899-514-2
 ; Sequence 2, Application US/08899514
 ; Sequence 2, Application US/08899514
 ; General Information:
 ; Applicant: HABUCHI, OSAMI
 ; Applicant: FUKUTA, MASAKAZU
 ; Title of Invention: POLYPEPTIDE OF GLYCOSAMINOGLYCAN
 ; Title of Invention: SULFOTRANSFERASE ORIGINATING FROM HUMAN AND DNA CODING
 ; Title of Invention: FOR THE POLYPEPTIDE
 ; Number of Sequences: 9
 ; Correspondence Address:
 ; Addressee: KNOBBE, MARTEENS, OLSON & BEAR, LLP
 ; Street: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
 ; City: NEWPORT BEACH
 ; State: CALIFORNIA
 ; Country: US
 ; ZIP: 92660
 ; Computer Readable Form:
 ; Medium Type: Floppy disk
 ; Computer: IBM PC compatible
 ; Operating System: PC-DOS/MS-DOS
 ; Software: PatentIn
 ; Current Application Data:
 ; Application Number: US/08/899,514
 ; Filing Date:
 ; Classification: 435
 ; Prior Application Data:
 ; Application Number:
 ; Attorney/Agent Information:
 ; Name: DANIEL E ALTMAN
 ; Registration Number: 34,115
 ; Reference/Document Number: TOYAM21.001AUS
 ; Telecommunication Information:
 ; Telephone: 714 760 0404
 ; Telefax: 714 760 9502
 ; Information for Seq ID No: 2:
 ; Sequence Characteristics:
 ; Length: 479
 ; Type: amino acid
 ; Topology: linear
 ; Molecule Type: protein
 ; US-08-899-514-2

Db 512 VNSPEEVKDLSTLL 526
 Qy 243 IRECVSHVYIAEATLKPFPFLGRYRIVRFEDLAREPLAEFRLAYAFTGLTLPQEA 302
 Db 349 MEVICNSMAKTLQTA-LQPDWLQGHYLIVRYEDLGVPVKTLRVEYDFVGLVSPMEQ 407
 Qy 303 WIHNITHGSG-IGKPIEAFTSSNARNAYSQMRHALPFTKLRVEYCAQALQLLGYRP 361
 Db 408 FALNMTSGSSSSRP--FVVSARNATQANANRALTTFQIQKVEEFCYQPMAVLYER 464
 Qy 362 VYSAQQDQLTLDLV 376
 Db 465 VNSPEEVKDLSTLL 479
 Qy 376 VYSAQQDQLTLDLV 376
 Db 479 VNSPEEVKDLSTLL 479
 Qy 479 VNSPEEVKDLSTLL 479
 Result 12
 US-19-949-016-6471
 ; Sequence 6471, Application US/0949016
 ; Patent No. 681239
 ; General Information:
 ; Applicant: VENTER, J. Craig et al.
 ; Title of Invention: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; File Reference: CLO01307
 ; Current Application Number: US/09/949,016
 ; Current Filing Date: 2000-04-14
 ; Prior Application Number: 60/241,755
 ; Prior Filing Date: 2000-10-20
 ; Prior Application Number: 60/237,768
 ; Prior Filing Date: 2000-10-03
 ; Prior Application Number: 60/231,498
 ; Prior Filing Date: 2000-09-08
 ; Number of SEQ ID Nos: 207012
 ; Software: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 6471
 ; LENGTH: 531
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-6471
 Query Match 32.0%; Score 656.5; DB 4; Length 531;
 Best Local Similarity 39.7%; Pred. No. 2.7e-63;
 Matches 149; Conservative 68; Mismatches 121; Indels 37; Gaps 12;
 Qy 34 SPAGGEDRTH-VLVLSWRSQSSFLQGFLQSPHPDVYLMPEPAHWWTLSQGSAATHM A 92
 Db 157 APEGVGDKRHMWYTFWRSQSSFLQGFLQSPHPDVYLMPEPAHWWTLSQGSAATHM A 92
 Qy 93 VRDLMRSIFLMDVFDAYMPQ---SRNLIS---AFFNWATSRLCSPPACASAPRGTSKQ 147
 Db 217 ARDMISALXRCDSLVPOLYSPAGSGGRNLITLGFGAATNKVTCSSPLCPARKEVGLV 276
 Qy 148 D-VCKTLCTROPSLAREACRSYSHVYKLEVRFFRNLOVLYPLSDPAINLRLVHLYRDP 205
 Db 277 DDRVCKK-CPPQLARFEECRKYRTLVIKVGRVFDVAVLPLRDPALDKVTHLYRDP 335
 Qy 206 RAVLRS-----REAGPILLARD-----NGILVLTGNGKRV-EADPH-LRL 242
 Db 336 RAVASSRTRSRHGLIRESQVRDRPRAHMPFLEAAGHKLGAKEVGPGDHYAIGA 395
 Qy 243 IRECVSHVYIAEATLKPFPFLGRYRIVRFEDLAREPLAEFRLAYAFTGLTLPQEA 302
 Db 396 MEVICNSMAKTLQTA-LQPDWLQGHYLIVRYEDLGVPVKTLRVEYDFVGLVSPMEQ 454
 Qy 303 WIHNITHGSG-IGKPIEAFTSSNARNAYSQMRHALPFTKLRVEYCAQALQLLGYRP 361
 Db 455 FALNMTSGSSSSRP--FVVSARNATQANANRALTTFQIQKVEEFCYQPMAVLYER 511
 Db 352 VYSAQQDQLTLDLV 376
 Qy

RESULT 14
 Qy 311 SGIGKPIEAFHTSSRNARNVSQAWRHALPFTKLRLRQEVOAGALQLLGYRPVYSADQORD 370
 Db 412 SGI-----YSTQRNSSEOFEKWRFSMPFKLADQVVAQCGPANRFLGYKLARDAAALTN 464
 Qy 371 LTLDLVLPKG 380
 Db 465 RSVSLLERG 474

RESULT 14
 US-0-655-878-2 Application US/08655878
 ; Sequence 2, Application US/08655878
 ; Patent No. 5827713
 ; GENERAL INFORMATION:
 ; APPLICANT: HABUCHI, OSAMI
 ; TITLE OF INVENTION: DNA CODING FOR SULFOTRANSFERASE
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE:
 ; STREET:
 ; CITY:
 ; STATE:
 ; COUNTRY:
 ; ZIP:
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC Compatible
 ; SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent.In
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/655,878
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME:
 ; REGISTRATION NUMBER:
 ; REFERENCE/DOCKET NUMBER:
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE:
 ; TELEFAX:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 458
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-655-878-2

Query Match 25.4%; Score 540.5; DB 2; Length 458;
 Best Local Similarity 34.3%; Pred. No. 1.3e-50; Mismatches 66; Indels 31; Gaps 8;
 Matches 124; Conservative

Qy 36 AGGEDRVRHVVLSSWRSQSSFLGOLFQSQHDFVYLMPEPAWHV--WTTLSQGSATLHNAY 93
 Db 107 AAPEPRHRVILMATTGTGSSFVGFFNQQCNIFLPEPLWHIEVTTFPGGANAVGSAI 166
 Qy 94 --RDLMRSFLCDMDVFDAYM-- PQSRNLIAAFFNWATSAFLCSPPACSAFPRTISKOD 148
 Db 167 VYRDLTQQLCDLYILESTISPAEEHTAALPQQSSHLSCBEPCTPSLKKVFEKYH 226
 Qy 149 VCKTLCTROPSLAREACRSYSHVYLKEFRFLNQVLYPLSDPALNRLVHVRDPRAY 208
 Db 227 CNRRCGPNTILAAECKRKHMLKTVTROLFQPLAEDPDLRIOLVDRDPRAY 286
 Qy 209 LRSREAAQGPILLARDNGIVGTGNGKVE-----ADPHRLJIREVCRSHVIAEATL 259
 Db 287 LVSRNVA-----FSSKEYSKWVAEGAPLQDEDFQRORGNCES-IRUSAEGJL 335
 Qy 260 KPPPFLRGYRLVRFEDLAREPLAEIRALYAFFGLTLPOLEAWHNIITHSGCIGKPIEA 319

RESULT 15
 US-0-9-015-188-2 Application US/09015188C
 ; Sequence 2, Application US/09015188C
 ; Patent No. 6399158
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, Kevin J
 ; APPLICANT: Tabas, Ira
 ; TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
 ; FILE REFERENCE: JEFF-0231
 ; CURRENT APPLICATION NUMBER: US/09/015,188C
 ; CURRENT FILING DATE: 1998-01-29
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 2

Query Match 25.0%; Score 513.5; DB 3; Length 411;
 Best Local Similarity 33.5%; Pred. No. 1.e-47; Mismatches 63; Indels 43; Gaps 12;
 Matches 139; Conservative

Qy 7 SSXTVYTILLA----QTTCCLLFLFISR---PGPSSPAGEDRV-----H 43
 Db 4 SWKAVLILALASTAQTAAIRTTAKSFHTCPG-LAEAGLAERUCEESPFTAYNLSRKTH 62

Query Match 44 VLVISSWRSQSSSPYQOLFSQHDPVYLMPEPAWHVWTL---SQGSA---ATLHMAVRD 95
 Db 63 ILILATTSRSQSSSPYQOLFSQHDPVYLMPEPAWHVQNTLIPRFQGKSPADRRVMLGASRD 122

Qy 96 LMRSFLCDMDVFDAYM--PQSRNLIAFFNWATSAFLCSPPACSAFPRTISKODVCK 151
 Db 123 LIRSPLYDCDLYLEPNYIKRPPVHNTTDRIFRGASRVLCSPVCDPPGADVLLEGDCV 182

Qy 152 TLTCTROPFLSAREACRSYSHVYLKEFRFLNQVLYPLSDPALNRLVHVRDPRAYLRS 211
 Db 183 RKCCDNLTVAASACRESHVARTKTVYPEVDLAVELPRNLKVQLVDRGFLAS 242

Qy 212 REAGPILLARDNGIVLGTGNGKWEADPHRLIREVCRSHVRIAATLKPPFLRGYRL 271
 Db 243 RSETFRDYTRMLWYGTGRKPYNLD--VTOLTITVCFDSNSNSVSTGMLR--PWLGKCKML 299

Qy 272 VRFEDLAREPLAEIRALYAFFGLTLPQLEAWHNIITG-GIGKPIEAFHTSSRNARYV 330
 Db 300 VRVEDLARNPMKTEETIIGFLGPLDSHVARYQNNTRGDPFLGK--HKYGTVRNSAAT 356

Qy 331 SOAWRHALPFTKLRVQEVCAAGQLLGYRPVYSADQORDLTLDVLPRGPDHFS 385
 Db 357 AEKWRFRUYDIVAFAQNAQCQVLAQUGKIAASBEELKNPSVSLVEBDFRPTS 411

Search completed: June 23, 2005, 08:52:11
 Job time : 15.6998 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 08:34:49 ; Search time 14.8883 Seconds
(without alignments)
1980.512 Million cell updates/sec

Title: US-10-697-828-7
Perfect score: 2057

Sequence: 1 MRLPRVSSTVMLSLLMVTG.....LPRGMDSFKWASSTEKQPE 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:

513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgm2_6_ptodata/1/iaa/5A_COMB.pep:
2: /cgm2_6_ptodata/1/iaa/5B_COMB.pep:
3: /cgm2_6_ptodata/1/iaa/5A_COMB.pep:
4: /cgm2_6_ptodata/1/iaa/5B_COMB.pep:
5: /cgm2_6_ptodata/1/iaa/5CTUS_COMB.pep:
6: /cgm2_6_ptodata/1/iaa/backfiles.pep:
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---|
| 1 | 1708 | 83.0 | 395 | 4 | US-09-949-016-7011 Sequence 7011, Ap |
| 2 | 1536.5 | 74.7 | 390 | 4 | US-09-949-016-5813 Sequence 6813, Ap |
| 3 | 1536.5 | 74.7 | 431 | 4 | US-09-949-016-0893 Sequence 8893, Ap |
| 4 | 1017.5 | 49.5 | 436 | 3 | US-09-045-284A-2 Sequence 2, Appli |
| 5 | 1017.5 | 49.5 | 386 | 3 | US-09-190-911-1 Sequence 1, Appli |
| 6 | 986.5 | 48.0 | 386 | 4 | US-09-786-240-11 Sequence 11, Appli |
| 7 | 647.5 | 31.5 | 484 | 4 | US-09-263-023-4 Sequence 4, Appli |
| 8 | 647.5 | 31.5 | 484 | 4 | US-09-167-4 Sequence 4, Appli |
| 9 | 647.5 | 31.5 | 531 | 4 | US-09-949-016-0471 Sequence 6471, Ap |
| 10 | 640 | 31.1 | 483 | 3 | US-09-263-023-2 Sequence 2, Appli |
| 11 | 636.5 | 30.9 | 608 | 4 | US-09-167-2 Sequence 9449, Ap |
| 12 | 549 | 26.7 | 479 | 3 | US-08-899-514-2 Sequence 2, Appli |
| 13 | 495 | 24.1 | 458 | 2 | US-08-655-078-2 Sequence 2, Appli |
| 14 | 96 | 4.7 | 468 | 4 | US-09-489-039A-14323 Sequence 14323, A |
| 15 | 91 | 4.4 | 487 | 4 | US-09-252-591A-23907 Sequence 23907, A |
| 16 | 90 | 4.4 | 933 | 1 | US-08-370-193A-8 Sequence 8, Appli |
| 17 | 90 | 4.4 | 933 | 1 | US-09-271-138A-9 Sequence 9, Appli |
| 18 | 90 | 4.4 | 933 | 4 | US-10-078-107-5 Sequence 5, Appli |
| 19 | 90 | 4.4 | 933 | 4 | US-10-077-151-5 Sequence 5, Appli |
| 20 | 90 | 4.4 | 933 | 4 | US-09-49-039A-12157 Sequence 12157, A |
| 21 | 90 | 4.4 | 294 | 4 | US-09-252-591A-18083 Sequence 18083, A |
| 22 | 89.5 | 4.4 | 471 | 4 | US-09-902-040-10244 Sequence 10244, A |
| 23 | 89.5 | 4.4 | 421 | 4 | US-09-818-780-17 Sequence 17, Appli |
| 24 | 88.5 | 4.3 | 268 | 4 | US-09-818-780-94 Sequence 94, Appli |
| 25 | 88.5 | 4.3 | 267 | 4 | US-09-252-591A-24244 Sequence 24244, A |
| 26 | 88.5 | 4.3 | 561 | 4 | US-09-252-591A-24244 Sequence 24244, A |

ALIGNMENTS

RESULT 1

US-09-949-016-7011
; Sequence 7011, Application US/09949016

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949-016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIORITY DATE: 2000-09-08
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO: 7011
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7011
Query Match 83.0%; Score 1708; DB 4; Length 395;
Best Local Similarity 83.0%; Pred. No. 1.4e-181;
Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0;
1 MRLPRESSSTVMLSLLMMQTGILVFLVSRQPSSPAGLGERVHVLSSWRSGSSFVGOLF
QY 1 MRLPRESSSTVMLSLLMMQTGILVFLVSRQPSSPAGLGERVHVLSSWRSGSSFVGOLF
Db 1 MWLPRVASSPTAVTALLAGTFLFLVSRGPSSPAGGERVHVLSSWRSGSSFVGOLF
Matchs 61 SQAPDVFYLMEPAWVWtLSSOSSAPALHMAVDLIRLSVFLCONDVFDAYLPQRNNSDL
QY 61 SQAPDVFYLMEPAWVWtLSSOSSAPALHMAVDLIRLSVFLCONDVFDAYLPQRNNSDL 120
Db 61 NQRFDVFLMEPAWVWtLSSQSSAA1HMAVRDVLVSVFLCDMDVFDAYLPQRNNSDL 120
121 FQAVASRALCSPPPVCEAPRGNTSESYCKPLGQAFCSSYSHVVLKEVRPFN 180
Db 121 FQAVASRALCSPPPVCEAPRGNTSESYCKPLGQAFCSSYSHVVLKEVRPFN 180
181 LOVLYPLISDPALNRLTHLVRDPRAVLRSEQTAKALDNGIVLGTGNTWYEAIDPLR 240
QY 181 LOVLYPLISDPALNRLTHLVRDPRAVLRSEQTAKALDNGIVLGTGNTWYEAIDPLR 240
Db 181 LOVLYPLISDPALNRLTHLVRDPRAVLRSEQTAKALDNGIVLGTGNTWYEAIDPLR 240
121 FQAVASRALCSPPPVCEAPRGNTSESYCKPLGQAFCSSYSHVVLKEVRPFN 180
QY 121 FQAVASRALCSPPPVCEAPRGNTSESYCKPLGQAFCSSYSHVVLKEVRPFN 180
241 VVNEVCRSHVRIAA1LKPPFLQDRYPLVRYBFLDAPLTVIRELAFTGSLTPQLE 300
Db 241 VVNEVCRSHVRIAA1LKPPFLQDRYPLVRYBFLDAPLTVIRELAFTGSLTPQLE 300
241 VVREVCVSHVRIAA1LKPPFLQDRYPLVRYBFLDAPLTVIRELAFTGSLTPQLE 300
QY 241 VVREVCVSHVRIAA1LKPPFLQDRYPLVRYBFLDAPLTVIRELAFTGSLTPQLE 300
301 TWIHNTHGSGPARRAFAKTSRDAISVSQAWRHTLPFAKIRRVQELGGALQLGYRS 360

RESULT 2

US-09-949-016-6813
; Sequence 8893, Application US/0949016
; Parent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 8893
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Human
; SEQ ID NO-949-016-8893

Query Match Score 1536.5; DB 4; Length 431;
Best Local Similarity 76.0%; Pred. No. 2.2e-162;
Matches 295; Conservative 30; Mismatches 62; Indels 1; Gaps 1;

Qy 1 MRLPRFSSVMLSLLMVQTG-ILVFLYSPRQVSPAGLGERVHVLSSWRSGSSSFVGQL 59
Db 42 MWLPRFSSVMLSLLMVQTG-ILVFLYSPRQVSPAGLGERVHVLSSWRSGSSSFVGQL 101

Qy 60 FSOHPDVFYLMEPAWHNTLTSQGSAAPALHMAVRDLRSVFLCDMDYFDYLPPRNISD 119
Db 102 FSQHPDVFYLMEPAWHNTLTSQGSAAPALHMAVRDLRSVFLCDMDYFDYLPPRNISD 161

Qy 120 LFQWAVSRALCSPPVCEAFAGNISSEVCPLCATPFGLAQEBASSSYHVLKEYRFF 179
Db 162 FWQWAVSRALCSPPVCEAFAGNISSEVCPLCATPFGLAQEBASSSYHVLKEYRFF 221

Qy 180 NLQVLYPLISDPALNLRVHVLDRPAVLRSREQTAKALARNDNGTIVLTGTNTWEDPRL 239
Db 222 NLQVLYPLISDPALNLRVHVLDRPAVLRSREQTAKALARNDNGTIVLTGTNTWEDPRL 281

Qy 240 RVYNECRSHYRIARALHKPPPELQDRYRLVRYEQLDPLTVIRELYAFTGLGLTPQL 299
Db 282 RLIREVCRSHYRIAEATLKKPPFLGRYRLVRFEDAREPIAETRAYAFTGLLTPQL 341

Qy 300 QTWHINITHGSGPGARAEAKTSRDALSQAMRHTLPFAKIRYQELCGALQOLGYR 359
Db 342 EAWHINTHGSIGKPKEAFTSSRURDVRLPFTKLRLQEVCA GALQOLGYR 401

Qy 360 SVHSELORDSLDPLIPRGMDSFKWAS 387
Db 402 PVYSDAQDQRLTDLVLPRGDHFMSMAS 429

RESULT 4

US-09-045-284A-2
; Sequence 2, Application US/09045284A
; Patent No. 6255192
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stephan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107US1
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
; SEQ ID NO-09-045-284A-2

Query Match Score 1017.5; DB 3; Length 386;
Best Local Similarity 56.4%; Pred. No. 1.5e-104;

RESULT 3

US-09-949-016-8893
; Sequence 8893, Application US/0949016
; Parent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

Matches 215; Conservative 43; Mismatches 112; Indels 11; Gaps 7;

Qy 1 MRLPRFSSTVNLMLMVOTGILVF--LVSROVPS-SPAGLGERVHVLVSSWSRGSSFPV 56
 Db 1 MLLPK--KMKLLFLYSGMALLPFRMHSNISLSMKAQPERMHVLVSSWSRGSSFPV 58

Qy 57 GOLFSDQHDFYLMEPAWHTWMTLSQGSAPALMAVRDLIRSVPFLCDMDYFDAYL-PWRR 115
 Db 59 GOLFSDQHDFYLMEPAWHTWMTLSQGSAPALMAVRDLIRSVPFLCDMSVFDAYMEPGPR 118

Qy 116 NISDLFOWAVERSALCSPPVCAPAFGNISSEBEVKPLCATRPFGLAQEASSYSHVILKE 175
 Db 119 RQSSLFOWERSALCSAPACDIIPODEIIPRAHCRLLCSCQOPPEVEKARSISHVILKE 178

Qy 176 VRFENIQVLYPLSDDPALNIRVHLVRDPAVLRSREQTAKALARNDGIVLGTNTW-B 234
 Db 179 VRFENIQVLYPLSDDPALNIRVHLVRDPAVLRSREQTAKALARNDGIVLGTNTW-B 238

Qy 235 ADPRLRVNVNECRSHVIRAAALHKPPFLQDRYLVRYEDLARDPLTVIRELYAFTGIG 294
 Db 239 EDQPYYVMQVICOSQLEIYK-TIQSLPKALQERYLTYK-SQWRHTLPKAIRRVQELCGGALQ 354

Qy 295 LTPOLOQTWINTHSGSGCARREPAKTSRDSLVSQWRHTLPKAIRRVQELCGGALQ 354
 Db 298 FLPHLQTWINHTTRKGGMG-DHAFTNARDALNSQMRWSLPLYEKVSRQLKACGDAMN 355

Qy 355 LLGYRSVHSLEORDLSLDL 375
 Db 356 LLGYRHVRSEQRNLDDL 376

RESULT 6

US-09-786-240-11

Sequence 11; Application US/09786240
 Patent No. 6558335

GENERAL INFORMATION:
 APPLICANT: INCUTIVE PHARMACEUTICALS, INC.
 ATTORNEY: TANG, Y. Tom
 CORLEY, Neil C.
 GOEGLER, Karl J.
 BAUGHN, Mariah R.
 LAL, Preeti
 YUE, Henry
 HILLMAN, Jennifer L.
 APPLICANT: AZIMZAI, Yalda

TITLE OF INVENTION: HUMAN TRANSPERASE PROTEINS
 FILE REFERENCE: PF-0592 PCT

CURRENT APPLICATION NUMBER: US/09/786 , 240
 CURRENT FILING DATE: 2002-03-12
 PRIORITY APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 09/113,642
 PRIORITY FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1999-05-11
 NUMBER OF SEQ ID NOS: 33

SEQ ID NO 11
 LENGTH: 386
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. 6558935 2617407CD1

US-09-190-911-1

Sequence 1; Application US/09190911
 Patent No. 6365365

GENERAL INFORMATION:
 APPLICANT: Bistrup, Annette
 ATTORNEY: Rosen, Steven D.
 APPLICANT: Tangemann, Kirsten
 APPLICANT: Hemmerich, Stefan
 TITLE OF INVENTION: GLYCOSTYL SULFOTRANSFBRASE -3
 FILE REFERENCE: 6510-107CIP
 CURRENT APPLICATION NUMBER: US/09/190,911
 EARLIER FILING DATE: 1998-11-12
 EARLIER APPLICATION NUMBER: 09/045,284
 EARLIER FILING DATE: 1998-03-20
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 386
 TYPE: PRT
 ORGANISM: H. sapiens

Query Match 49.5%; Score 1017.5; DB 3; Length 386;
 Best Local Similarity 56.1%; Pred. No. 1.5e-104;
 Matches 215; Conservative 43; Mismatches 112; Indels 11; Gaps 7;

Qy 1 MRLPRFSSTVNLMLMVOTGILVF--LVSROVPS-SPAGLGERVHVLVSSWSRGSSFPV 56
 Db 1 MLLPK--KMKLLFLYSGMALLPFRMHSNISLSMKAQPERMHVLVSSWSRGSSFPV 58

Qy 57 GOLFSDQHDFYLMEPAWHTWMTLSQGSAPALMAVRDLIRSVPFLCDMDYFDAYL-PWRR 115
 Db 59 GOLFSDQHDFYLMEPAWHTWMTLSQGSAPALMAVRDLIRSVPFLCDMSVFDAYMEPGPR 118

Qy 116 NISDLFOWAVERSALCSPPVCAPAFGNISSEBEVKPLCATRPFGLAQEASSYSHVILKE 175
 Db 119 RQSSLFOWERSALCSAPACDIIPODEIIPRAHCRLLCSCQOPPEVEKARSISHVILKE 178

Qy 176 VRFENIQVLYPLSDDPALNIRVHLVRDPAVLRSREQTAKALARNDGIVLGTNTW-B 234
 Db 179 VRFENIQVLYPLSDDPALNIRVHLVRDPAVLRSREQTAKALARNDGIVLGTNTW-B 238

Qy 235 ADPRLRVNVNECRSHVIRAAALHKPPFLQDRYLVRYEDLARDPLTVIRELYAFTGIG 294
 Db 239 EDQPYYVMQVICOSQLEIYK-TIQSLPKALQERYLTYK-SQWRHTLPKAIRRVQELCGGALQ 297

Qy 295 LTPOLOQTWINTHSGSGCARREPAKTSRDSLVSQWRHTLPKAIRRVQELCGGALQ 354
 Db 298 FLPHLQTWINHTTRKGGMG-DHAFTNARDALNSQMRWSLPLYEKVSRQLKACGDAMN 355

Qy 355 LLGYRSVHSLEORDLSLDL 375
 Db 356 LLGYRHVRSEQRNLDDL 376

RESULT 5

US-09-786-240-11

Sequence 1; Application US/09786240
 Patent No. 6558335

GENERAL INFORMATION:
 APPLICANT: INCUTIVE PHARMACEUTICALS, INC.
 ATTORNEY: TANG, Y. Tom
 CORLEY, Neil C.
 GOEGLER, Karl J.
 BAUGHN, Mariah R.
 LAL, Preeti
 YUE, Henry
 HILLMAN, Jennifer L.
 APPLICANT: AZIMZAI, Yalda

TITLE OF INVENTION: HUMAN TRANSPERASE PROTEINS
 FILE REFERENCE: PF-0592 PCT

CURRENT APPLICATION NUMBER: US/09/786 , 240
 CURRENT FILING DATE: 2002-03-12
 PRIORITY APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 09/113,642
 PRIORITY FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1999-05-11
 NUMBER OF SEQ ID NOS: 33

SEQ ID NO 11
 LENGTH: 386
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: Incyte ID No. 6558935 2617407CD1

US-09-786-240-11

Query Match 48.0%; Score 986.5; DB 4; Length 386;
 Best Local Similarity 56.3%; Pred. No. 4.2e-101;
 Matches 215; Conservative 40; Mismatches 114; Indels 13; Gaps 9;

Qy 1 MRLPRFSSTVNLMLMVOTGILVF--LVSROVPS-SPAGLGERVHVLVSSWSRGSSFPV 56
 Db 1 MLLPK--KMKLLFLYSGMALLPFRMHSNISLSMKAQPERMHVLVSSWSRGSSFPV 58

Qy 57 GOLFSDQHDFYLMEPAWHTWMTLSQGSAPALMAVRDLIRSVPFLCDMDYFDAYL-PWRR 115
 Db 59 GOLFSDQHDFYLMEPAWHTWMTLSQGSAPALMAVRDLIRSVPFLCDMSVFDAYMEPGPR 118

Qy 116 NISDLFOWAVERSALCSPPVCAPAFGNISSEBEVKPLCATRPFGLAQEASSYSHVILKE 175
 Db 119 RQSSLFOWERSALCSAPACDIIPODEIIPRAHCRLLCSCQOPPEVEKARSISHVILKE 178

Qy 176 VRFENIQVLYPLSDDPALNIRVHLVRDPAVLRSREQTAKALARNDGIVLGTNTW-B 234
 Db 179 VRFENIQVLYPLSDDPALNIRVHLVRDPAVLRSREQTAKALARNDGIVLGTNTW-B 238

Qy 235 ADPRLRVNVNECRSHVIRAAALHKPPFLQDRYLVRYEDLARDPLTVIRELYAFTGIG 294
 Db 239 EDQPYYVMQVICOSQLEIYK-TIQSLPKALQERYLTYK-SQWRHTLPKAIRRVQELCGGALQ 297

Qy 295 LTPOLOQTWINTHSGSGCARREPAKTSRDSLVSQWRHTLPKAIRRVQELCGGALQ 354
 Db 298 FLPHLQTWINHTTRKGGMG-DHAFTNARDALNSQMRWSLPLYEKVSRQLKACGDAMN 355

Qy 355 LLGYRSVHSLEORDLSLDL 375
 Db 356 LLGYRHVRSEQRNLDDL 376

NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 6471
 LENGTH: 531
 TYPE: PRT
 ORGANISM: Human
 US-09-949-016-6471

Query Match 31.5%; Score 647.5; DB 4; Length 531;
 Best Local Similarity 36.8%; Pred. No. 4.-e-63; Gaps 9;
 Matches 140; Conservative 76; Mismatches 117; Indels 47;

QY :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 33 SPAGLGERH-VIQLSSRGSSESSPVGOLFISQHDVFTIMEPAWHDWTLSQSAPALHMA 91
 QD :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 157 APECGVDKRRHMVTFITRSGSSEPFEGELNQNPEVFVFLYEPMVHWQKLYPGDAVSLOGA 216

PY :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 92 VRDLIRSVLFCDMDVFDAYLP---WWRNISDL--FONAVSRALCSPPVCEAFARG--NIS 144
 QD :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 217 ARDMLSALYRCDSVFLQYSQVCPAYKEVGLV 276

QY :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 145 SEEVCKPLCATRPGLAQACSYSSVHIVLKEYRFFNLQVLYPLSDPALNLRIVHLVRDP 204
 QD :|||:|||:|||:|||:|||:|||:|||:|||:|||:
 277 DDRVCKK-CPPQRALARFEECRKYRTLVIGRVVEDAVALPLRDPLDKVTHLVRDP 335

PY :|||:|||:|||:|||:|||:|||:|||:|||:
 205 RAVLRSRECTAKALARDNCIVLGCTNGTWEADPR-----
 QD :|||:|||:|||:|||:|||:|||:
 336 RAVASSRIRSRRHGLIRESSQVRSR-----DERAHRMMPF 375

QY :|||:|||:|||:|||:|||:|||:
 239 --LRRVNEVCRSHVRIAALHKPPPQLQDRYRLVRYEDLARDPLTVIRELYAFTCLGL 295
 QD :|||:|||:|||:|||:|||:|||:|||:
 390 YHALGAMEVCSMALKTLQTAL-QPDDWNOQGHLYLUVVREDLVPVKTLLRYYDFVGILV 448

QY :|||:|||:|||:|||:|||:|||:|||:
 296 TPOLQTWIHNITHGSGPGRARAEFKTSRDLASVSOAMRHTLFPFKIRRQELCGGALQL 355
 QD :|||:|||:|||:|||:|||:|||:|||:
 449 SPENQFANMTSSEGGSSSK--PPVVSANATOAQANAVRALTQFQIKOVEEBCYQPMVA 506

PY :|||:|||:|||:|||:|||:|||:
 356 LGYRSVHSSELEQDLSLDL 375
 QD :|||:|||:|||:|||:|||:
 507 LGYERVNNSPBEVKDLSKTL 526

RESULT 11
 US-09-471-867-2 ; Sequence 2, Application US/09471867
 ; FILE REFERENCE: TOYAMA1.00IAUS
 ; CURRENT APPLICATION NUMBER: US/09/471,867
 ; PATENT NO. 645289
 ; GENERAL INFORMATION:
 ; APPLICANT: Uchimura, Kenji
 ; APPLICANT: Muramatsu, Hideki
 ; APPLICANT: Kadomatsu, Kenji
 ; APPLICANT: Kannagi, Reiji
 ; APPLICANT: Habuchi, Osami
 ; APPLICANT: Muramatsu, Takashi
 ; TITLE OF INVENTION: POLYPEPTIDE OF N-ACTYGLUCOSAMINE-6-O-SULFOTRANSFERASE
 ; TITLE OF INVENTION: DNA ENCODING THE SAME
 ; FILE REFERENCE: TOYAMA1.00IAUS
 ; CURRENT APPLICATION NUMBER: US/09/471,867
 ; CURRENT FILING DATE: 1999-12-23
 ; EARLIER APPLICATION NUMBER: US 09/263,023
 ; EARLIER FILING DATE: 1999-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-54007
 ; EARLIER FILING DATE: 1998-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-177844
 ; EARLIER FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FabSEQ for Windows Version 3.0
 ; SEQ ID NO 2 ; Score: 640; DB 4; Length 483;
 ; LENGTH: 483
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-263-023-2 ; Query Match 31.1%; Score 640; DB 3; Length 483;
 ; Best Local Similarity 36.2%; Pred. No. 2.8e-62;
 ; Matches 145; Conservative 73; MisMatches 1.6; Indels 64; Gaps 10;
 ; SEQ ID NO 2 ; Score: 640; DB 4; Length 483;
 ; LENGTH: 483
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-263-023-2 ; Query Match 31.1%; Score 640; DB 3; Length 483;
 ; Best Local Similarity 36.2%; Pred. No. 2.8e-62;
 ; Matches 145; Conservative 73; MisMatches 1.6; Indels 64; Gaps 10;

Db 209 KVQSSPLCPAYRKEVGLIDRVCXK-CPPOLARFEECRKYRTVVKGVRVEDVAVL 267
 185 YPLLSDPALNLRIVHLYRDPRAVLRSREQTAKALARDNGIVLGTNTWVEADPR---- 238
 Qy 268 APLKQDPLDLKIVHLYRDPRAVASSRIRSRLRHRESLOVRSR----DPRAHRMPP 321
 Db 239 -----LRYVNNECRSHVRIAELAHKPPFLQDRYRLVRYED 275
 Qy 322 LEAAGHKLGAKEGMGGPADYHALGAMEVICNSMAKTLQL-QPPDWLQGHYLVRYED 380
 Db 276 LARDPLTIVRLYEAFTGLGLTQLOTHNITHGSGPAREFAFKITSRDSLVSQWRH 335
 Qy 381 LVGDPVKTRLRYDFGLLVSSEMEQPALNTMSGSSSK--PFVVSARNATQAANAWRT 438
 Db 336 TLPFKAFTKRRVQBLGCGALQJLGYRSVHSELFFORDLSLDL 375
 Qy 439 ALTFQQIKQVEEFCYQMAVIGYERVSNSPEEVKDLSKTLL 478
 Db 439 ALTFQQIKQVEEFCYQMAVIGYERVSNSPEEVKDLSKTLL 478

RESULT 12
 US-09-949-016-9449
 ; Sequence 9449, Application US/09949016
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C10/01307
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 9449
 ; LENGTH: 608
 ; TYPE: PRT
 ; ORGANISM: Human
 ;
 US-09-949-016-9449

Query Match 30.9%; Score 636.5; DB 4; Length 608;
 Best Local Similarity 37.1%; Pred. No. 9.8e-67; Mismatches 117; Indels 47; Gaps 9;

Db 95 LIRSVFLCDMDVFDAYLP---MWNNTSDL--FQWAVSRLASPPVCEAFARGN -NISSEE 147
 Qy 297 MLSALYRCDSFLQYSPAGSGRNLTTLGFGAATMKVUVSSPLCPAYREVVGLVDDR 356
 Db 148 YCKPLCATRPGLAQACSSYSHVVLKEVRFENLQVLYPLSDPAINLRLTVHLVRDRAV 207
 Qy 357 VCKK-CPPQRALARFEECRKRTLVKGVRVFDVAVLAPLDRDPLDKVHLVRDRAV 415
 Db 208 LRSRECTAKALARDNGIVLGTNTWVEADPR----- 238
 Qy 416 ASRIRHRGILRESLOVRSR----DPRAHRMPPLEAAGHKLGAKKEGVGPADYHA 469
 Db 239 LRVVNNECRSHVRIAELAHKPPFLQDRYRLVRYEDLARDPLTIVRELYAFTGLGLTQ 298
 Qy 470 LGAMEVICNSMAKTLQL-QPPDWLQGHYLVRYEDLARDPLTIVRELYAFTGLGLTQ 298
 Db 299 LQWVINITRGSGPAREFAFKITSRDSL/SQAWRHTLPPFKIIRRQELCGALQQLGY 358
 Qy 529 MEQFALNMTSGGSSSK--PFVVSARNATQAANAWRTALFOQIKQVEEFCYQPMAVLGY 586
 Db 351 SAEGLG-LQPAWLRGRMLVRYEDVARGPLQKAREMPFAGIPLTPQVEDWIKNTQAH 409
 359 RSVHSELEQRDLSDLI 375

Db 587 ERVNSPEEVKDLSKTLL 603
 Db RESULT 13
 US-09-999-514-2
 ; Sequence 2, Application US/08999514
 ; Patent No. 5910581
 ; GENERAL INFORMATION:
 ; APPLICANT: HABUCHI, OSAMI
 ; TITLE OF INVENTION: POLYPEPTIDE OF GLYCOSAMINOGLYCAN
 ; TITLE OF INVENTION: SULFOTRANSFERASE ORIGINATING FROM HUMAN AND DNA CODING
 ; TITLE OF INVENTION: FOR THE POLYPEPTIDE
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: KNOBBE, MARTENS, OLSON & BEAR, LLP
 ; STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
 ; CITY: NEWPORT BEACH
 ; STATE: CALIFORNIA
 ; COUNTRY: US
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/899,514
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIORITY APPLICATION DATA:
 ; PRIOR APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DANIEL E ALTMAN
 ; REGISTRATION NUMBER: 34,115
 ; REFERENCE/DOCKET NUMBER: TOYAM21.001AUS
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 714 760 0404
 ; TELEFAX: 714 760 9502
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 479
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ;
 US-08-999-514-2

Query Match 27.3%; Score 561.5; DB 2; Length 479;
 Best Local Similarity 38.6%; Prod. No. 1.6e-53;
 Matches 142; Conservative 55; Mismatches 136; Indels 35; Gaps 11;

Qy 31 PSSPAGGGERVHVLSWRSGSFVGOLFQSHQPDVYVLMPEAWHWDTLS--QGSAPAL 88
 Db 123 PRPAPAVGPRRHVLLMATTGTTSSFVGEFFNQGNIFVLFEPWHERTVSFRGAAJA 182
 Qy 89 HMVAV--RDLIRSYFLCMDVFDAY---LPWRRNISLDFQWAVSRLASPPVCEAFARGN 143
 Db 183 GSALVYRDBVLQKFLCDIYVLERFITIPEDHTQFMRGGSRSCLDDPVCFPKV-KV 241
 Qy 144 SSEEVCK----PLCATRPFGLAQACSSYSHVVLKEYRFFENLQVLYPLSDPAINLRLV 198
 Db 242 FEKYHCNRRCGPLNT----LAAECRKRKEMLKAVRIOLEFLOPLAEDPRDLRVI 297
 Qy 199 HLVRDPRAVLRSR-----EQPATKALARDNGIVLGTNTWVEADPRURVNEVCRSHVR 251
 Db 298 QLVRDPRAVLRASRMVAFAGKYTWKWLDE---GDQGLRREEVQLR---GNCESIRL 350
 Qy 252 IAEEALKHPPLQDRLYRLVRELYEDLARDPLTIVRELYAFTGLGLTQLOQTWHNTHGSG 311
 Db 351 SAEGLG-LQPAWLRGRMLVRYEDVARGPLQKAREMPFAGIPLTPQVEDWIKNTQAH 409

Qy 312 PGARREAFKITSRDALSVSQAMWRHTLPPAKIRVOELCGGALQLLGYSVHSELEQRDLS 371
 Db 410 DGS--GIYTTQKNSEQQFKWRSMPFKLAQVVAQPCPAMRFLGYKLARDAAALTTRS 466

Qy 372 LDLLPRG 379
 Db 467 VSLLEBERG 474

RESULT 14

; Sequence 2, Application US/09015188C
 ; Patent No. 6399318
 ; GENERAL INFORMATION:
 ; APPLICANT: Tabas, Ira
 ; TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
 ; 6-Sulfotransferase
 ; FILE REFERENCE: JEFF-0231
 ; CURRENT APPLICATION NUMBER: US/09/015,188C
 ; CURRENT FILING DATE: 1998-01-29
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 2
 ; LENGTH: 411
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-015-188-2

Query Match 26.7%; Score 549; DB 3; Length 411;
 Best Local Similarity 36.4%; Pred. No. 3e-32;
 Matches 136; Conservative 58; Mismatches 160; Indels 20; Gaps 9;

Query Match 24.1%; Score 495; DB 2; Length 458;
 Best Local Similarity 36.1%; Pred. No. 3.e-46;
 Matches 129; Conservative 59; Mismatches 125; Indels 44; Gaps 13;

Qy 25 LVSIQVPSSA--GLGERHVLVLSMWSGSSFVGOLFQSHPDVYFVYMBPAWHVWDTL- 80
 Db 42 LAERLCEEESTFAYNLSRKTHILATTRGSSPFVGOLFQNHLQHLDVFYLFBPLHVQNTLI 101

Qy 81 --SQGAPS---LHMAYVLDLRSVFLCMDFDVAL--PWRNISDLFQAVASRALC 130
 Db 102 PRFTQGKSPDRRVMNLGASDILRLSRYDCOLYFLENYTKPPVNHTTDRFGASRVIC 161

Qy 131 SPPVCEAFAFGNISSEE--VKCKPLCATRPFELAQEACSSYSHVILKEVRFFNLQVLYPLIS 189
 Db 162 SRPVCDPPGADLVLEGGDVKRCGCLLNLTVAACRERSHVAIKTVRVPEVNDLRAVVE 221

Qy 190 DPAINLRIVLVLDRDAVISREQTAKALARDNGVLTGNTWYEADPRLRVNEVCISH 249
 Db 222 DPRNLKVQLVDRDGTLSLSRSETFRDTRLRMLWGTGKPKYNLD--VTOJLTVCEBF 279

Qy 250 VRIAAAHLKPPFLQDRYFLVRYEDLARPLVTIRELTAFTGGLTPQQTWHNHTIG 309
 Db 280 SNSVSTGLMR-PPMJKGVNMLVRYEDLARNPMKCTTEIYGLFLG1PLDSHVARVNQNNTRG 338

Qy 310 SGPGARAREFKITSRDALSVSQAMWRHTLPPAKIRVOELCGGALQLLGYSVHSELEQRDLS 369
 Db 339 D-PTLGHKHKGTV-RNSAATDAEKMRFLRSYDIVAFQAQQVLAQGLGYKTIASEEELKN 396

Qy 370 LSDLJLPRGMDSF 383
 Db 397 PSVSLVVEERFRP 410

RESULT 15

; Sequence 2, Application US/08655878
 ; Patent No. 5827713
 ; GENERAL INFORMATION:
 ; APPLICANT: FUKUTA, MASAKAZU
 ; APPLICANT: HABUCHI, OSAMI
 ; TITLE OF INVENTION: DNA CODING FOR SULFOTRANSFERASE
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE:

Search completed: June 23, 2005, 08:52:10
 Job time : 16.8883 secs

THIS PAGE BLANK (USPTO)

| | | |
|--|--|---|
| Copyright (c) 1993 - 2005 Compugen Ltd. | GenCore version 5.1.6 | Aab95367 Human pro Adc18590 Human sof Aau69414 Lung small |
| OM protein - protein search, using SW model | | Adr14283 Human NF-Ab668582 Drosophila |
| Run on: June 23, 2005, 08:32:43 ; Search time 64.4109 Seconds (without alignments) | 3578.727 Million cell updates/sec | Aay73919 Mouse gly Aay72640 Human gly Abp81554 Human cor Aae15438 Human dru Adi11086 Novel hum Adi61235 Human tyr Abn85237 Human pro Abx1139 Human DIT |
| Title: US-10-697-828-9 | Perfect score: 3147 | Aaw72863 Glycosami Aay72639 Human gly Abb1556 Human int Abn85236 Mouse pro Abb1557 Mouse int Aay72638 Mouse gly Aau11275 Murine in |
| Sequence: 1 MNRNYGAMMDWDAHYKMFN.....LIENICWTLMDRGYPKFMD 596 | Scoring table: BL0SUMM2 Gapext 0.5 | |
| Searched: 2105692 seqs, 386760381 residues | Total number of hits satisfying chosen parameters: 2105692 | |
| Minimum DB seq length: 0 | Maximum DB seq length: 2000000000 | |
| Post-processing: Minimum Match 0* | Maximum Match 100% | |
| | Listing first 45 summaries | |
| Database : A_Geneseq_16Dec04:* | | RESULT 1 ID AAY72641 AC XX XX DT 02-MAY-2001 (first entry) XX DB Human glycosyl sulfotransferase-6 (GST-6) fragment. XX KW Human; glycosyl sulfotransferase-6; GST-6; immunosuppressive; therapy; KW selectin binding inhibitor; gene therapy; inflammation; KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes; KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis; KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; anaemia; KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia; KW denyeinating disease; cirrhosis; ulcerative colitis; allergic rhinitis; KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis; KW asthma; hypersensitivity; rheumatic fever; tissue rejection. XX OS Homo sapiens. XX PN WO200106015-A1. XX PD 25-JAN-2001. XX PF 19-JUL-2000; 2000WO-US019741. XX PR 20-JUL-1999; 99US-0144694P. PR 13-JUN-2000; 2000US-00593828. XX DR (RESC) UNIV CALIFORNIA. PA PI Roben SD, Lee JK, Hemmerich S; XX DR N-PSDB; AAD02702, AAD02703. |
| | | SUMMARIES |
| Result No. | Score | Query Match Length DB ID Description |
| 1 | 3147 | 100.0 596 4 AAY72641 Human gly |
| 2 | 3147 | 100.0 1222 4 AAY72642 Human gly |
| 3 | 3142 | 99.8 1212 6 AAE33541 Human nov |
| 4 | 3142 | 99.8 1222 6 ABU11849 Human sec |
| 5 | 2835 | 90.1 1207 4 AAY72643 Mouse gly |
| 6 | 2835 | 90.1 1207 6 AAE33542 Human nov |
| 7 | 1821 | 57.9 755 4 AAB93735 Human pro |
| 8 | 1821 | 57.9 755 8 ADP56673 Human pro |
| 9 | 920 | 29.2 480 4 AAB94410 Human pro |
| 10 | 637 | 20.2 128 4 ADG27796 Human nov. |
| 11 | 11 | 19.5 314 5 AAB93826 Human pro |
| 12 | 568 | 18.0 125 8 ADJ12217 Human sec |
| 13 | 568 | 18.0 126 2 AAY14415 Human sec |
| 14 | 147 | 4.7 410 7 ADN95522 Human BBC |
| 15 | 147 | 4.7 411 2 AAW61100 Keratan s |
| 16 | 147 | 4.7 411 5 AAE25556 Human cho |
| 17 | 147 | 4.7 411 6 ABO03503 Angiogene |
| 18 | 147 | 4.7 411 7 ADJ68589 Human hea |
| 19 | 147 | 4.7 411 8 DQD18725 Human sof |
| 20 | 136 | 4.3 474 3 AAB34722 Human sec |
| 21 | 136 | 4.3 625 8 ADR1287 Human pro |
| 22 | 136 | 4.3 958 3 AAY51120 Human SAR |
| 23 | 136 | 4.3 958 8 ADD83102 Human PRO |
| 24 | 134 | 4.3 483 2 AAY31656 Mouse N-a |
| 25 | 133 | 4.2 484 2 AAY31657 Human N-a |

| Sequence 596 AA: | | | | | | | |
|------------------|--|--------|-----------|----------|------------|----|------|
| | Query Match | Score | DB 4; | Length | 596; | | |
| Q | Best Local Similarity | 100.0* | Pred. No. | 1.6e-30; | Mismatches | 0; | Gaps |
| Q | Matches 596; Conservative | 0; | Indels | 0; | | | |
| b | 1 MNRYNGAMMDWDAHYKMFWDIHGNSPMASIQEAQAEFKWRWTQFVNVTQFQESTETIT 60 | | | | | | |
| b | 1 MNRYNGAMMDWDAHYKMFWDIHGNSPMASIQEAQAEFKWRWTQFVNVTQFQESTETIT 60 | | | | | | |
| b | 61 RIAVFVGPYINVSSCRFDSSNPGLOISLNVNTNEHVSIVTVDYHNLTTRFNLYLGFGFP 124 | | | | | | |
| b | 61 RIAVFVGPYINVSSCRFDSSNPGLOISLNVNTNEHVSIVTVDYHNLTTRFNLYLGFGFP 124 | | | | | | |
| b | 121 ASVADQGQITRGLGTQAIKPVYRDRRIFFPGEFKENIAGVLILCISVLTVLTFQWRFLYS 188 | | | | | | |
| b | 121 ASVADQGQITRGLGTQAIKPVYRDRRIFFPGEFKENIAGVLICISVLTVLTFQWRFLYS 188 | | | | | | |
| b | 181 FRKLMMRMLILVLTALWFTELLDWSTCSQPICAWTRTEAEGSKSLSSEGHNLDLPDVV 241 | | | | | | |
| b | 181 FRKLMMRMLILVLTALWFTELLDWSTCSQPICAWTRTEAEGSKSLSSEGHNLDLPDVV 241 | | | | | | |
| b | 241 ITSLPGSGAEIILKQLFFNNSDFLYIRVPTAYIDIPETELEIDSFTVDACEWKVSDIRSCHF 301 | | | | | | |
| b | 241 ITSLPGSGAEIILKQLFFNNSDFLYIRVPTAYIDIPETELEIDSFTVDACEWKVSDIRSCHF 301 | | | | | | |
| b | 301 RLLRGWQLSQLVDTDKLHLQNLNHLHEPNRGKLACYPAMNKDKCKFKRRESLPEQSOMKG 366 | | | | | | |
| b | 301 RLLRGWQLSQLVDTDKLHLQNLNHLHEPNRGKLACYPAMNKDKCKFKRRESLPEQSOMKG 366 | | | | | | |
| b | 361 AFDRDABYIARLRRHLTYPSARPVLSSGKSMPLKQFFQEYVGLASMRALYTVDPRAW 424 | | | | | | |
| b | 361 AFDRDABYIARLRRHLTYPSARPVLSSGKSMPLKQFFQEYVGLASMRALYTVDPRAW 424 | | | | | | |
| b | 421 IYSMLYNSPKSLYSLKNVPEHLAKLFKTEGGKCKCNLNSGYAFYEPLRKELSKSKSNAY 480 | | | | | | |
| b | 421 IYSMLYNSPKSLYSLKNVPEHLAKLFKTEGGKCKCNLNSGYAFYEPLRKELSKSKSNAY 480 | | | | | | |
| b | 481 SLLSHMLANTAAALRINTDLPLTSYOLVKFEDVHPQKTTTERIFATFGIPISPLASLNQ 541 | | | | | | |
| b | 481 SLLSHMLANTAAALRINTDLPLTSYOLVKFEDVHPQKTTTERIFATFGIPISPLASLNQ 541 | | | | | | |
| b | 541 ILFATSTNLFLPYGEGETSPNTNWKONLPRDEIKLIEINICWTLMDRLGYPKMDF 596 | | | | | | |
| b | 541 ILFATSTNLFLPYGEGETSPNTNWKONLPRDEIKLIEINICWTLMDRLGYPKMDF 596 | | | | | | |

| | | | |
|---|--------------|---------------|---|
| glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis; Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia; demyelininating disease; cirrhosis; ulcerative colitis; allergic rhinitis; myocarditis; adult respiratory distress syndrome; eczema; psoriasis; asthma; hypersensitivity; rheumatic fever; tissue rejection. | Homo sapiens | Key Domain | Location/Qualifiers 851..1123 /label=C-terminal_sulfotransferase_domain |
|---|--------------|---------------|---|

25-JAN-2001. 19-JUL-2000; 2000WO-US019741.

20-JUL-1999; 99US-0144694P.

13-JUN-2000; 2000US-00593828.

(RESCC) UNIV CALIFORNIA.

Rosen SD, Lee JK, Hemmerich S;

WPI, 2001-134471/14.

N-PSDB; AAD0702, AAD02704.

New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications.

Claim 3: Fig 5B; 128pp; English.

The present sequence is human glycosyl sulfotransferase-6 (GST-6). GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non-sulphated selectin ligand. GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in inhibiting a selectin mediated binding event. GST is useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis nodosa, polyosmotic dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's disease, Grave's disease, adrenatitis, hypoparathyroidism, pernicious anaemia, demyelinating diseases, cirrhosis, ulcerative colitis, dematitis, myocarditis, regional enteritis, ulcerative respiratory distress syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation

Sequence 1222 AA;

| | Query Match | Match | Score | DB | Length |
|----|--|---------|---------------------|-----------|--------|
| 1 | MNRYNGAMMDWDAHYKMFWDHGNSPMASQEAQAEFKRKWTOFVNNTFQMESTIT | 100.0% | 3147; | DB 4; | 1222; |
| 2 | MNRYNGAMMDWDAHYKMFWDHGNSPMASQEAQAEFKRKWTOFVNNTFQMESTIT | 100.0%; | Pred. No. 4.9e-107; | | |
| 3 | 627 MNRYNGAMMDWDAHYKMFWDHGNSPMASQEAQAEFKRKWTOFVNNTFQMESTIT | | Mismatches 0; | Indels 0; | Gaps 0 |
| 4 | 61 RIATVFGPYINVSRCFDSSNPGLQISLNVNNTTEHVSVIYTDYHNLKTRFNYLGFFG | | | | 60 |
| 5 | 687 RIATVFGPYINVSRCFDSSNPGLQISLNVNNTTEHVSVIYTDYHNLKTRFNYLGFFG | | | | 686 |
| 6 | 121 ASVADQGQTTRFGGTQAIKVPRHDRIIFPGPKENIAGVHLICSLVLTQFRFLYS | | | | 120 |
| 7 | 747 ASVADQGQTTRFGGTQAIKVPRHDRIIFPGPKENIAGVHLICSLVLTQFRFLYS | | | | 746 |
| 8 | 181 FRKLMRWLILVIALWFELDDWVSTCSQPIAKWTRTEAEGSKSLSUSEGHMQLPDVV | | | | 240 |
| 9 | 807 FRKLMRWLILVIALWFELDDWVSTCSQPIAKWTRTEAESSKSLSUSEGHMQLPDVV | | | | 866 |
| 10 | 241 ITSLPGSGAEILKQLFNNSSDFLYIRVPTAYIDIPETELEIDSVDACEKMVSDIRSGH | | | | 300 |

| | | | | |
|---|---|---|------|---|
| Db | 972 | PFDRAEYIRALRRHVVYPSARPVLSSGSWTLKLHFFQEVLGTSMRALYIVRDPRAW | 1031 | CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95993 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention |
| Oy | 421 | IYSMLYNKSPSLSYLKNEYPEHLAKLKFTEGGKGKCNLNSGYAEEPEPLRKELSKSKSNAV | 480 | CC XX |
| Db | 1032 | IYSVLYGSPSLSYLKNEYPEHLAKLKFKEGRSKCNSGYSAYEYESLKKELEBISQNAI | 1091 | CC Sequence 755 AA; |
| Qy | 481 | SILSHIWLAATAAALRINTDLPPTSYOLYKPEEDIHVPPQTKTERIAFAGIPSPASINQ | 540 | Query Match 57.9%; Score 1821; DB 4; Length 755; |
| Db | 1092 | SILSHIWLAATAAALRINTDLPPTSYOLYKPEEDIHVPPQTKTERIAFAGIPSPASINQ | 1151 | Best Local Similarity 99.4%; Pred. No. 1..1e-173; |
| Qy | 541 | ILPFATSTNLFLYLPYEGEISPTNTNWKNLPROBEIKLLENICWTLMRQLYKPEFMD | 596 | Matches 342; Conservative 1; Mismatches 1; Indels 0; Gaps 0; |
| Db | 1152 | ILPFATSTNLFLYLPYEGEISPTNTNWKTNLPRDEIKLLENICWTLMRQLYKPEFMD | 1207 | Qy 1 MNRYNGAMMDYDAHYKMFWDHGNSPMASIOEAQAAEFKKRWTQFVNVTQFQMBESTIT 60 |
| Db | 411 | MNRYNGAMMDYDAHYKMFWDHGNSPMASIOEAQAAEFKKRWTQFVNVTQFQMBESTIT | 470 | Db 411 MNRYNGAMMDYDAHYKMFWDHGNSPMASIOEAQAAEFKKRWTQFVNVTQFQMBESTIT 470 |
| RESULT 7 | | | | Qy 61 RIAYVFGPYINVSSCRFPIDSSNPGLQISLNVNNTTEHVSIVTDYHNLTKEPNYLGGGF 120 |
| AAB93735 | | | | Db 471 RIAYVFGPYINVSSCRFPIDSSNPGLQISLNVNNTTEHVSIVTDYHNLTKEPNYLGGGF 530 |
| ID | | | | Qy 121 ASVADQGQTTRFLGQTAIVKPVHRDIIIPPGFKENIAVGLLICSLSVLTFQWREYLS 180 |
| XX | | | | Db 531 ASVADQGQTTRFLGQTAIVKPVHRDIIIPPGFKENIAVGLLICSLSVLTFQWREYLS 590 |
| AC | | | | Qy 181 FRKLMLRMLVILVIALWFIELDWWSQSOPICAKWTRTEAGSKKSISSEGHHMDLPDVV 240 |
| XX | | | | Db 591 FRKLMLRMLVILVIALWFIELDWWSQSOPICAKWTRTEAGSKKSISSEGHHMDLPDVV 650 |
| DT | 26-JUN-2001 | (first entry) | | Qy 241 ITSLPQSGAEILKQLPFNSSDFLYIPTVTAIDIPETELEIDSFVDACEWKVSDIERSGHF 300 |
| XX | | | | Db 651 ITSLPQSGAEILKQLPFNSSDFLYIPTVTAIDIPETELEIDSFVDACEWKVSDIERSGHF 710 |
| DE | | | | Qy 301 RLLRGWQSLVQDTKHLQNLHLPNRGKLAQYFAMNKDKKKR 344 |
| XX | | | | Db 711 RLLRGWQSLVQDTKHLQNLHLPNRGKLAQYFAMNKDKKKR 754 |
| Human protein sequence SEQ ID NO:13382. | | | | RESULT 8 |
| XX | | | | ADP56673 standard; Protein: 755 AA. |
| DE | | | | ID ADP56673 |
| XX | | | | AC ADP56673; |
| KW | | | | XX DT 12-AUG-2004 (first entry) |
| Human; primer; detection; diagnosis; antisense therapy; gene therapy. | | | | XX Human protein which is a chondroitin 6 sulphotransferase 3 homologue. |
| XX | | | | DE |
| OS | | | | XX cancer detection; large intestine; oesophagus; stomach; lungs; pancreas; |
| Homo sapiens. | | | | XX liver; colon; human; chondroitin 6 sulphotransferase 3; CH6T3. |
| XX | | | | XX Homo sapiens. |
| EP1074617-A2. | | | | XX PN JP2004147505-A. |
| XX | | | | XX PD 27-MAY-2004. |
| PD | 07-FEB-2001. | | | XX PR 28-OCT-2002; 2002JP-00312927. |
| XX | | | | XX PA (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO. |
| PF | 28-JUL-2000; | 20000EP-00116126. | | XX PA (FJRE) FUJIREBIO KK. |
| PR | 29-JUL-1999; | 99JP-00248036. | | XX PA (SEGK) SEIKAGAKU KOGYO CO LTD. |
| PR | 27-AUG-1999; | 99JP-00300253. | | XX DR WPI: 2004-434531/41. |
| PR | 11-JAN-2000; | 20000JP-00118776. | | XX N-PSDB; ADP56673. |
| PR | 02-MAY-2000; | 20000JP-00183767. | | XX |
| PR | 09-JUN-2000; | 20000JP-00241839. | | XX Novel nucleic acid such as DNA, useful as probe for identifying or detecting cancer tissue of esophagus, stomach, lungs, pancreas, liver, kidney or colon. Preferably large intestine tissue. |
| PA | (HELI-) HELIX RES INST. | | | XX Example 1; SEQ ID NO 2; 24pp; Japanese. |
| XX | | | | XX The invention relates to a novel nucleic acid having 40-1000 base pairs |
| PI | Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J; | | | CC |
| PI | Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T; | | | CC |
| XX | | | | CC |
| DR | 2001-318749/34. | | | CC |
| XX | | | | CC |
| PS | Claim 8; SEQ ID NO 13382; 2537pp + Sequence Listing; English. | | | CC |
| XX | | | | CC |
| CC | The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. | | | CC |
| CC | Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. | | | CC |
| CC | Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification. Where a primer set comprises: | | | CC |
| CC | (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the nucleotide sequence comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the 5'-end sequence of the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and | | | CC |

XX Human; primer; detection; diagnosis antisense therapy; gene therapy.
 XX
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-00116126.
 XX PR 29-JUL-1999; 99JP-00248036.
 XX PR 27-AUG-1999; 99JP-0030053.
 XX PR 11-JAN-2000; 2000JP-0011876.
 XX PR 02-MAY-2000; 2000JP-00183767.
 XX PR 09-JUN-2000; 2000JP-00241999.
 XX PA (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsubi T;
 XX DR WPI; 2001-318749/34.
 XX PR Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX PS SEQ ID NO 13638; 2537pp + Sequence Listing; English.
 XX CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH1363 to AAH18742 represent human cDNA sequences; AAH95446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.
 XX SQ Sequence 314 AA;
 Query Match 19.5%; Score 615; DB 4; Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.1e-52;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 470 KELSKSKSNAVSLSHMLNTAAALRINTDLPSTSQLVKFEDVHPQQKTERIAPFL 529
 Db 2 KNYQPKNSNAVSLSHMLNTAAALRINTDLPSTSQLVKFEDVHPQQKTERIAPFL 61
 Qy 530 GIPUSASINQILFATSTNLFLYPIEGEISPTINTVWQNLPDEIKIENICWTLMDRL 589
 Db 62 GIPUSASINQILFATSTNLFLYPIEGEISPTINTVWQNLPDEIKIENICWTLMDRL 121
 Qy 590 GYPKFM 596
 Db 122 GYPKFM 128
 RESULT 11
 AAB93826 26-JUN-2001 (first entry)
 ID AAB93826 standard; protein; 314 AA.
 XX AC AAB93826;
 XX DT ADJ12217 standard; protein; 125 AA.
 XX DE Human protein sequence SEQ ID NO:13638.

AC ADJ12217; DT 20-MAY-2004 (first entry)

XX Human secreted protein SeqID 71.

KW human; secreted; cancer; haematopoietic disease; anaemia; multiple myeloma; reproductive system disorder; prostatitis; insipinal hernia; musculoskeletal disease; systemic lupus erythematosus; gout; cardiovascular disease; arrhythmia; hypernatraemia; fetal disease; fetal alcohol syndrome; Down's syndrome; excretory disease; urinary incontinence; renal disorder; neural; sensory disease; Alzheimer's disease; meningitis; respiratory disease; emphysema; glomerulonephritis; digestive disease; diabetes; diarrhoea; irritable bowel syndrome; epithelial disease; scleroderma; epidermolysis bullosa; cytotoxic; antianemic; antiarthritic; antiallergic; anti-HIV; immunosuppressive; antiinflammatory; antipsoriatic; antibacterial; osteopathic; dermatological; antigout; immunomodulator; antiarthritis; nephrotropic; uroprotective; anabolic; anabolic, hypertensive and antidiabetic; anabolic; antidiabetic; tranquilizer; antiparkinsonian; neuroprotective; antiparkinsonian; tranquilizer; antidiabetic; anabolic; hypertensive; vulnerary.

OS Homo sapiens.

PN US2004010132-A1.

XX PD 15-JAN-2004.

XX PF 30-OCT-2001; 2001US-00984429.

XX PR 09-OCT-1997; 97US-0061463P.

PR 09-OCT-1997; 97US-0061527P.

PR 09-OCT-1997; 97US-0061529P.

PR 09-OCT-1997; 97US-0061532P.

PR 09-OCT-1997; 97US-0061536P.

PR 09-OCT-1997; 97US-0071498P.

PR 08-OCT-1998; 98WO-US021142.

PR 08-APR-1999; 99US-00288143.

PR 01-NOV-2000; 2000US-0244591P.

XX PA (ROSE/) ROSEN C A.

PA (BREW/) BREWER L A.

PA (DUAN/) DUAN R D.

PA (RUBE/) RUBEN S M.

PA (FLOR/) FLORENCE K A.

PA (GREE/) GREENE J A.

PA (YOUN/) YOUNG P E.

PA (FEAR/) FERRIE A M.

PA (FLOR/) FLORENCE C.

PA (EBNE/) EBNER R.

PA (OLSE/) OLSEN H.

XX PI Rosan CA, Brewer LA, Duan RD, Ruben SM, Florence KA, Greene JM; Young PE, Ferrie AM, Yu G, Florence C, Ebner R, Olsen H;

XX WPI: 2004-090518/09.

DR N-PSDB; ADU12161.

XX PT New isolated nucleic acids and polypeptides, useful for diagnosing, treating, preventing or ameliorating diseases or disorders e.g. cancer, anaemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's disease.

XX PS Claim 11; SEQ ID NO 71; 28pp; English.

XX PT This invention relates to novel polynucleotides encoding human secreted proteins. Specifically, it refers to the vectors, host cells, recombinant and synthetic methods for producing human polynucleotides, polypeptides and antibodies. Furthermore, it relates to screening methods to identify agonists and antagonists that can be used to inhibit or enhance the production and function of the secreted proteins. The present invention

CC describes these compositions as useful for diagnosing, treating or preventing disorders such as cancer, haematopoietic diseases including anaemia and multiple myeloma, reproductive system disorders including prostatitis and insipinal hernia, musculoskeletal diseases including systemic lupus erythematosus and gout, cardiovascular disease including arrhythmia and hypernatraemia, mixed fetal diseases including fetal alcohol syndrome and Down's syndrome, excretory diseases including urinary incontinence and renal disorders, neural or sensory disease including Alzheimer's disease and maningitis, respiratory disease including emphysema and occupational lung disease, endocrine diseases including diabetes and glomerulonephritis, digestive diseases including portal hypertension and irritable bowel syndrome and connective tissue or epithelial diseases including scleroderma and epidermolysis bullosa. As such, there are various activities such as cytotoxic, anti-anemic, antiarthritic, antiasthmatic, anti-HIV, immunosuppressive, antiinflammatory, antipsoriatic, antibacterial, osteopathic, cardiotonic, dermatological, antigen, immunomodulator, antiarrhythmic, cardiant, nootropic, antilipemic, nephrotropic, uropathic, neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive and vulnery. This polypeptide is a human secreted protein of the invention. NOTE: This sequence does not appear in the printed specification but has been obtained in electronic format from the US patent office at the following web site www.seqdata.uspto.gov/sequence.html; Document ID: 2004010132.

XX SQ Sequence 125 AA;

Query Match 18.0%; DB 8; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 MRWILILVIALWFIELDWSTCSQPIKAKWTRTEAGSKKSLSSEGHHMDLPDVVITSL 244
Db 1 MRWILILVIALWFIELDWSTCSQPIKAKWTRTEAGSKKSLSSEGHHMDLPDVVITSL 60

Qy 245 PGSSAEILKQLQFFNSDFFLYIRVPTAYDIPETELEIDSFVDAECKW 291
Db 61 PGSSAEILKQLQFFNSDFFLYIRVPTAYDIPETELEIDSFVDAECKW 107

RESULT 13
AAV14415
ID AAV14415 standard; protein; 126 AA.
XX AC AAV14415;
AC AAV14415;
XX DT 17-AUG-1999 (first entry)
XX DE Human secreted protein encoded by gene 5 clone HSABG21.
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy; diagnosis; tissue; tumour; neurodegenerative disorder; leukaemia; developmental abnormality; foetal deficiency; blood; allergy; renal; immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma; inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS; cognitive disorder; schizophrenia; prostate; obesity; osteosoclast; thymus; osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion; endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX OS Homo sapiens.
XX PN WO9919339-A1.
XX PD 22-APR-1999.
XX XX 98WO-US021142.
XX PF 08-OCT-1998;
XX PR 09-OCT-1997;
PR 09-OCT-1997;
PR 09-OCT-1997;
PR 09-OCT-1997;
PR 09-OCT-1997;
PR 09-OCT-1997;
PR 09-OCT-1997;

XX (HUMA-) HUMAN GENOME SCI INC.
 PA XX WPI; 2003-876899/81.
 PI XX DR N-PSDB; ADN95522.
 PI XX Example 1; SEQ ID NO 445; 176pp; English.
 DR XX This invention relates to a method of differentially modulating the growth or differentiation of blood endothelial cells (BEC) or lymphatic endothelial cells (LEC) comprising contacting endothelial cells with a composition comprising an agent that differentially modulates blood or lymphatic endothelial cells. Treating hereditary lymphoedema comprises identifying a human subject with lymphoedema and with a mutation in at least one allele of a gene encoding a LEC protein, where the mutation correlates with lymphoedema in human subjects, and with the proviso that the LEC protein is not VEGF-C, and administering to the subject a composition comprising a lymphatic growth agent selected from VEGF-C or VEGF-D polypeptides and polynucleotides. The invention may be useful for the development of compounds with an antiangiogenic, cytostatic, vasotropic or antiinflammatory activity or for gene therapy. The method is useful in modulating the growth or differentiation of blood endothelial cells or lymphatic endothelial cells, in treating hereditary lymphoedema, in screening for an endothelial cell disorder or predisposition to the disorder or in monitoring the efficacy or toxicity of a drug on endothelial cells. The agent is useful in manufacturing a medicament for the differential modulation of blood vessel endothelial cell or lymphatic vessel endothelial cell growth or differentiation. The lymphatic growth agent may also be used in manufacturing a medicament for the treatment of hereditary lymphoedema resulting from a mutation in a LEC gene or of other diseases involving the lymphatic vessels, such as various inflammatory diseases and cancer metastasis via the lymphatic system. The present sequence is that of a human LEC/BEC differentially expressed protein which is related to the method of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer using the source data given in table 14 of the specification.

XX
 PS XX New isolated human genes and the secreted polypeptides they encode.
 PS XX Claim 11; Page 186; 226pp; English.
 PS XX This sequence represents a secreted human protein encoded by the gene clone detailed in the descriptor line. The gene can be used to generate fusion proteins by linking to the gene to a human immunoglobulin Fc portion (e.g. AAX19002) for increasing the stability of the fused protein as compared to the human protein only. The invention relates to 53 novel genes and their fragments (nucleic acid sequences: AAX7901-X79064; amino acid sequences AAY1441-Y14464) which are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Also pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the 53 polynucleotides, based on which tissues they are most highly expressed (see AAX7901 for described uses)

SQ Sequence 126 AA;

Query Match 18.0%; Score 568; DB 2; Length 126;
 Best Local Similarity 100.0%; Pred. No. 1.5e-48;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 185 MRWLLILVITALWFIELDDWSTCSQPICAKWTEAEQSKKSLSSEGHMQLPDVVITSL 244
 Db 1 PGSSAELIKQQLFFNNSSDFLYIRVTPAYDIPETELEIDSFDACEWK 291
 Qy 61 PGSSAELIKQQLFFNNSSDFLYIRVTPAYDIPETELEIDSFDACEWK 107

RESULT 14
 ID ADN95522 standard; protein: 410 AA.
 AC XX ADN95522;
 AC XX DT 01-JUL-2004 (first entry)
 DE Human BEC/LEC-related protein sequence SeqID445.
 XX growth; differentiation; blood endothelial cell; BEC;
 KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
 KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
 KW vasotropics; antiinflammatory; gene therapy; endothelial cell disorder;
 KW inflammatory disease; cancer metastasis; lymphatic system; human.
 XX Homo sapiens.
 OS XX
 PN XX WO2003080640-A1.
 PD XX 02-OCT-2003.
 XX
 PF XX 07-MAR-2003; 2003WO-US006900.
 PR XX 07-MAR-2002; 2002US-0363019P.
 PA XX (LUDW-) LUDWIG INST. CANCER RES.
 PA (LICN) LICENTIA LTD.
 Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
 PI XX
 QY XX

DR WPI; 2003-876899/81.
 XX N-PSDB; ADN95522.
 PS XX This invention relates to a method of differentially modulating the growth or differentiation of blood endothelial cells (BEC) or lymphatic endothelial cells (LEC) comprising contacting endothelial cells with a composition comprising an agent that differentially modulates blood or lymphatic endothelial cells. Treating hereditary lymphoedema comprises identifying a human subject with lymphoedema and with a mutation in at least one allele of a gene encoding a LEC protein, where the mutation correlates with lymphoedema in human subjects, and with the proviso that the LEC protein is not VEGF-C, and administering to the subject a composition comprising a lymphatic growth agent selected from VEGF-C or VEGF-D polypeptides and polynucleotides. The invention may be useful for the development of compounds with an antiangiogenic, cytostatic, vasotropics or antiinflammatory activity or for gene therapy. The method is useful in modulating the growth or differentiation of blood endothelial cells or lymphatic endothelial cells, in treating hereditary lymphoedema, in screening for an endothelial cell disorder or predisposition to the disorder or in monitoring the efficacy or toxicity of a drug on endothelial cells. The agent is useful in manufacturing a medicament for the differential modulation of blood vessel endothelial cell or lymphatic vessel endothelial cell growth or differentiation. The lymphatic growth agent may also be used in manufacturing a medicament for the treatment of hereditary lymphoedema resulting from a mutation in a LEC gene or of other diseases involving the lymphatic vessels, such as various inflammatory diseases and cancer metastasis via the lymphatic system. The present sequence is that of a human LEC/BEC differentially expressed protein which is related to the method of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer using the source data given in table 14 of the specification.

XX
 SQ Sequence 410 AA;

Query Match 4.7%; Score 147; DB 7; Length 410;
 Best Local Similarity 18.2%; Pred. No. 2.6e-59;
 Matches 89; Conservatve 58; Mismatches 149; Indels 194; Gaps 16;

Oy 187 WTLILVIALWFIELDDWSTCSQPICAKWTEAEQSKKSLSSEGHMQLPDVVITSL 229
 Db 5 WKA VLLILALASIAIQYTAIRPTAKSFTHTCPGLAEGLAERUC---EESPTFAVNL SR 59

Db 230 EGHHMDLPDVVITSLPGSGAELIKQQLFFNNSSDFLYIRVTPAYDIPETELEIDSFDACEWK 289
 Oy 60 KTH---ILLATRPGSSFVGQLFNQHLDVFLPEPL-----93

Db 290 WKVDISRGHFPRLRGWLQSLVQDTKHLQNLTHBPNRQGKLAQYPAMNNDKCKRKFKRRE 349
 Oy 94 -----YHVQNTLIPRFTOG-----KSPADRRV 115

Oy 350 SLEPQRSMQKGAFDRAEYI-----RARRHLVYPPSARPV-----LS 387
 Db 116 MGASRMLRSILSYDCDLYFLENYIKPPVNHTTDR1FRRGASRVLCSPVCDPPGPADLV 175

Db 388 LSSGSWTLKLHFOEYVGA-----SMRALYIVRD 416
 Oy 176 LEEGDCVRKCGLLNLTVAAACRSRSHVAIKTVRPDVENDLRLAVLPDLKVLQVLRD 235

Db 417 PRAWIYSLYNNSKPSLVSLSKVNPEHLAKLFIEGGKGKCNLNSGYAFYEPLARKELS SK 476
 Db 236 PRGTLASRSETFRDTY-----RLMLWYGR-----KPNLIDVTL-TIVQEDF 279

Oy 477 SNAVS---LLSHMLANTAAIRINTDLP-----RPLRQKXTTERIFAPLGIPPL- 533
 Db 280 SNSVSTCLMRPFWLKG-----KMLRYEDLARNPMKKTTEEYGFGLTPLD 325

Oy 534 -----SPASINQILFATSNLFYLPYEGISPTNTVWWQNLPDEIKLLEN 581
 Db 326 SHVARWQNTRGDPTLGKHYGTVRN-----SAATAEKWRFLSYDIVAFQNA 375

QY 582 CWTMLDRLG Y 591

Db 376 CQVLAQLG_r 385
 RESULT 15
 AAW61100 standard; protein; 411 AA.
 ID AAW61100
 XX
 AC AAW61100;
 XX DT 14-SEP-1998 (First entry)
 XX DE Keratan sulphate 6-sulphotransferase.
 XX KW KSGal6ST; keratan sulphate 6-sulphotransferase;
 KW chick chondroitin 6-sulphotransferase; C6ST; phage Lambda; hybridization;
 KW expression vector; COS-7 cells; Bluescript plasmid; Galactose;
 KW keratan sulphate; chondroitin.
 XX OS Homo sapiens.
 XX PN EP845533-A2.
 XX PD 03-JUN-1998.
 XX PF 27-NOV-1997; 97EPB-00309564.
 XX PR 29-NOV-1996; 96JPP-00320535.
 XX PA (SEIKAGAKU KOGYO CO LTD.
 XX PI Fukuta M, Habuchi O;
 DR WPI, 1998-288750/26.
 DR N-PDB; AAV3618.
 PT Keratan sulphate 6-sulpho-transferase polypeptide - transfers sulphate
 PT from sulphate donor to galactose 6-hydroxy group etc.
 XX PS Claim 4; Page 15-16; 21PP; English.
 XX CC This sequence produces the protein KSGal6ST (Keratan sulphate 6-
 CC sulphotransferase), which has a molecular weight of about 46700. The cDNA
 CC sequence of this protein was obtained by radiolabelling the cDNA of chick
 CC chondroitin 6-sulphotransferase (C6ST) and using this as a probe in a
 CC random oligonucleotide-primed labelling method. Human foetal brain cDNA
 CC was inserted into a phage Lambda gt11 cloning vector whereby the clones
 CC containing the KSGal6ST were obtained by hybridization using the prepared
 CC probe. The positive clones were subcloned into a recombinant expression
 CC vector and used to transform COS-7 cells, from which cells expressing
 CC KSGal6ST can be selected. The phage cDNA inserts were isolated and
 CC subcloned into a Bluescript plasmid. Deletion clones were then prepared
 CC from which both strands were sequenced by the Sanger method. The KSGal6ST
 CC of the invention transfers the sulphate from a sulphate donor to
 CC galactose 6-OH groups in keratan sulphate, but does not transfer sulphate
 CC to chondroitin, chondroitin sulphate A or C, dermatan sulphate or CDNS
 CC
 XX SQ Sequence 411 AA;

Query Match 4.7%; Score 147; DB 2; Length 411;
 Best Local Similarity 18.2%; Pred. No. 2.7e-05;
 Matches 89; Conservative 58; Mismatches 149; Indels 194; Gaps 16;

Qy 187 WIILVIALWPIEL-----LDWWSTC-----SOPICAKWTRTEAEGSKKSLS 229
 5 WKAVVLLALLASIAIQYTAIRTFAKSFHTCPGLABAGLRLC---EBSPTFAYNLSR 59

Db 230 EGHHMPLPDVYITSPLPGSAEILKOLFFNSSDFLYTRPVAYIDIPETELEIDSFVDACE 289

Qy 60 KTH-----IILILATRSGSSFVGOLFQHQDVFVLFEP----- 93

Qy 290 WKVSDIRSGHFRLLRGWLQSIVQDPTKLHQNIIHLHEPNRGKLAQYFAMNKDKKRKFKRE 349

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 08:34:49 ; Search time 22.4643 Seconds
(without alignments)

Title: US-10-697-828-9

Perfect score: 3147

Sequence: 1 MNRYNGAMMDWDAHYKMFN.....LIENICWTLMDRGYPKFMD 596

Scoring table: BL0SUMM2

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
 1: /cgri2_6/prodata/1/iaa/5a_COMB.pep:
 2: /cgri2_6/prodata/1/iaa/5b_COMB.pep:
 3: /cgri2_6/prodata/1/iaa/6a_COMB.pep:
 4: /cgri2_6/prodata/1/iaa/6b_COMB.pep:
 5: /cgri2_6/prodata/1/iaa/PCTUS_COMB.pep:
 6: /cgri2_6/prodata/1/iaa/backfiles/pep:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Score | Match Length | DB ID | Description |
|------------|-------------|--------------|-------|----------------------|
| 1 | 568 | 18.0 | 126 | US-09-288-143-71 |
| 2 | 147 | 4.7 | 411 | US-09-015-188-2 |
| 3 | 134 | 4.3 | 483 | US-09-263-023-2 |
| 4 | 134 | 4.3 | 483 | US-09-471-067-2 |
| 5 | 133 | 4.2 | 484 | US-09-263-023-4 |
| 6 | 133 | 4.2 | 484 | US-09-471-067-4 |
| 7 | 133 | 4.2 | 531 | US-09-949-016-6471 |
| 8 | 133 | 4.2 | 608 | US-09-949-016-9449 |
| 9 | 123.5 | 3.9 | 395 | US-09-949-016-7011 |
| 10 | 119.5 | 3.8 | 479 | US-08-899-014-2 |
| 11 | 119.5 | 3.8 | 390 | US-09-949-016-6813 |
| 12 | 119.5 | 3.8 | 431 | US-09-949-016-8893 |
| 13 | 115.5 | 3.7 | 386 | US-09-786-040-11 |
| 14 | 108.5 | 3.4 | 386 | US-09-045-084A-2 |
| 15 | 108.5 | 3.4 | 386 | US-09-190-911-1 |
| 16 | 108 | 3.4 | 826 | US-09-252-991A-22143 |
| 17 | 103.5 | 3.3 | 324 | US-09-270-067-4224 |
| 18 | 103 | 3.3 | 582 | US-09-091-025-17 |
| 19 | 101.5 | 3.2 | 594 | US-08-187-151-2 |
| 20 | 101 | 3.2 | 650 | US-09-248-096A-16387 |
| 21 | 100 | 3.2 | 1060 | US-09-248-096A-15476 |
| 22 | 99.5 | 3.2 | 601 | US-09-252-991A-22143 |
| 23 | 99 | 3.1 | 376 | US-09-150-141-7 |
| 24 | 99 | 3.1 | 376 | US-09-174-193-7 |
| 25 | 99 | 3.1 | 376 | US-09-374-024-7 |
| 26 | 99 | 3.1 | 376 | US-09-374-492-7 |
| 27 | 99 | 3.1 | 376 | US-09-107-133-7 |

ALIGNMENTS

RESULT 1
US-09-288-143-71
; Sequence 71, Appli
; Patent No. 6433339
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: PZ018P1
; CURRENT APPLICATION NUMBER: US/09-288,143
; CURRENT FILING DATE: 1993-04-08
; EARLIER APPLICATION NUMBER: PCT/US98/21142
; EARLIER FILING DATE: 1998-10-08
; EARLIER APPLICATION NUMBER: 60/061,463
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,529
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/071,498
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 2,19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 71
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: SITE
; NAME/KEY: SITE
; LOCATION: (126)
; OTHER INFORMATION: xaa equals stop translation
US-09-288-143-71

Query Match Score 568; DB 4; Length 126;
Best Local Similarity 100.0%; Pred. No. 2e-53; Indels 0; Gaps 0;

QY 185 MRWLILVIALWFELDWSTCSQPICKAKWTRTEAGSKKSLSSEGHHMDLPDVITSL 244
Db 1 MRWLILVIALWFELDWSTCSQPICKAKWTRTEAGSKKSLSSEGHHMDLPDVITSL 60

QY 245 PGSGAEBILQLQFFNSDFLYIPTYRATYDIPETELEIDSFDACEWK 291
Db 61 PGSGAEBILQLQFFNSDFLYIPTYRATYDIPETELEIDSFDACEWK 107

RESULT 2

US-09-015-188-2
; Sequence 2, Application US/09015188C
; Patent No. 6399358
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J
; TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
; TITLE OF INVENTION: 6-Sulfotransferase
; FILE REFERENCE: JEFF-0231
; CURRENT APPLICATION NUMBER: US/09/015,188C
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-015-188-2

Query Match Score 147; DB 3; Length 411;
Best Local Similarity 1.8-2%; Pred. No. 9, 6e-07; Indels 194; Gaps 16;
Matches 89; Conservative 58; Mismatches 149; Score 134; DB 3; Length 483;
; Pred. No. 3.3e-05;
; Mismatches 81; Indels 32; Gaps 9;

Qy 187 WILIVLIALWIEL-----LDWSTIC-----SOPICAKWTRTEAGSKSKLSS 229
Db 5 WKAVILLALASIAIQYTAIRFTAKSFHTCPGLAEAGLAERLC-----EESPTFAVNLSR 59
Qy 230 EGHMMDLPDVITSLPGSAAELKQLKFNSSDLYIRVPTAYDIPETELEDSFVIDACE 289
Db 60 KTH-----LILLATRSCSSFSVQFLNQHDFVFL-----93
Qy 290 WKVSDIRSSHFRLLRGWLOSVOTKLHLQNTNLHEPNRKLLAQYFAMNKDKKRKFRRRE 349
Db 94 -----YHQTNLIPRFTQG-----KSPADRRV 115
Qy 350 SLPORSONKGAFDRAEYI-----RALRRHLVYYPSSARPV-----LS 387
Db 116 MLGASDRLLRSLYCDLYFLENYIKPPPVNHTTDRIFRGASRVLCSRPVCDPPGPADLV 175
Qy 388 LSSGSWTLKHFDEVLG-----SMRALYIVRD 416
Db 176 LEEGSDCVRKCGLNLTVAEACERSHVAIKTVRVEYNNDLRLAVERPLNLKVQLVRD 235
Qy 417 PRATYSMLYNSKRESLYSLRNVPBEHLAKF1KIEKGKGNCKNLNSGYAFFEYEPKELRSK 476
Db 236 PRGLASRSETFRDTY-----RWRWLYGTGR ---KPYNLDVDTQL-TTVCEDP 279
Qy 477 SNAVS -L-LSHMLPLANTAAALARINTDLIPLTSYOLVKPFDIVHPPQKTERIFPAFLGIPL- 533
Db 280 SNSVSTGLNMRPPWKG-----KMLVRYEDLARNPMKCCTEYFGFLGIPLD 325
Qy 534 -----SPASLNQILFATSTNLFLYPIFGEISPTNTVWVKONLPRDEIKLHENI 581
Db 326 SHYARWIQINNTRGDPTLGRKYGTVRN-----SAATAEKWRPLSYDIVAFQNA 375
Qy 582 CWTLMDRUY 591
Db 376 CQVLAQLOGY 385

RESULT 3
; Sequence 2, Application US/09263023
; Patent No. 6037159
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; TITLE OF INVENTION: DNA ENCODING THE SAME

Query Match Score 134; DB 4; Length 483;
Best Local Similarity 26.2%; Pred. No. 3.3e-05;
Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 408 MRALYIVRDPRAWIYSMLYNSKPSLY--SLKNVPE----HIAKLFKIEGGK-GKCNLNS 459
Db 278 LKVIHLVDPRA-VASSRIRSRHGGLRESLQVRSRDPRAHMPFLEAAGHKLGAKKEGM 336
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-263-023-2

Query Match Score 134; DB 3; Length 483;
Best Local Similarity 26.2%; Pred. No. 3.3e-05;
Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 408 MRALYIVRDPRAWIYSMLYNSKPSLY--SLKNVPE----HIAKLFKIEGGK-GKCNLNS 459
Db 278 LKVIHLVDPRA-VASSRIRSRHGGLRESLQVRSRDPRAHMPFLEAAGHKLGAKKEGM 336
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-263-023-2

Query Match Score 134; DB 3; Length 483;
Best Local Similarity 26.2%; Pred. No. 3.3e-05;
Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 408 MRALYIVRDPRAWIYSMLYNSKPSLY--SLKNVPE----HIAKLFKIEGGK-GKCNLNS 459
Db 278 LKVIHLVDPRA-VASSRIRSRHGGLRESLQVRSRDPRAHMPFLEAAGHKLGAKKEGM 336
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-471-867-2

Query Match Score 134; DB 4; Length 483;
Best Local Similarity 26.2%; Pred. No. 3.3e-05;
Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 408 MRALYIVRDPRAWIYSMLYNSKPSLY--SLKNVPE----HIAKLFKIEGGK-GKCNLNS 459
Db 278 LKVIHLVDPRA-VASSRIRSRHGGLRESLQVRSRDPRAHMPFLEAAGHKLGAKKEGM 336
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-471-867-2

Qy 460 GYAFYEPLRKELSKSKSNAVSLSHLWANTAAALRINTDLPLTSYOLVKFEDIVHFPQ 519
 Db 337 GGPDYHAL-----GAMEVICNSMAKTLQTAQ-PDDWLOGHLYLVRVEDLVGDPV 386
 Qy 520 KTTERIFAGIPLSPA---SINOILPATSTNLFLYLYGEISPTN----TNWKONLP 571
 Db 387 KTLRVDYDFGLVSPMEQFALNMTCGSSSKPFPV----VSARNATOAAANAWRTALT 441
 Qy 572 RDEIKLIENICWTMLDRGPK 593
 Db 442 FQQIKQVEEFCYQPMNAVLYER 463

RESULT 5
 US-09-263-023-4
 ; Sequence 4, Application US/09263023
 ; Patent No. 6037159
 ; GENERAL INFORMATION:
 ; APPLICANT: Uchimura, Kenji
 ; APPLICANT: Muramatsu, Hideki
 ; APPLICANT: Kadomatsu, Kenji
 ; APPLICANT: Kannagi, Reiji
 ; APPLICANT: Habuchi, Osami
 ; APPLICANT: Muramatsu, Takashi
 ; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYGLUCOSAMINE-6-O-SULFOTRANSFERASE AND FILE REFERENCE: TOYAM41.00IAUS
 ; CURRENT APPLICATION NUMBER: US/09/263,023
 ; CURRENT FILING DATE: 1999-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-54007
 ; EARLIER FILING DATE: 1998-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-177844
 ; EARLIER FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 4
 ; LENGTH: 484
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-471-867-4

; APPLICANT: Muramatsu, Takashi
 ; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYGLUCOSAMINE-6-O-SULFOTRANSFERASE AND FILE REFERENCE: TOYAM41.00IAUS
 ; CURRENT APPLICATION NUMBER: US/09/471,867
 ; CURRENT FILING DATE: 1999-12-23
 ; EARLIER APPLICATION NUMBER: US 09/263,023
 ; EARLIER FILING DATE: 1999-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-54007
 ; EARLIER FILING DATE: 1998-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-177844
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 4
 ; LENGTH: 484
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-471-867-4

Query Match 4.2%; Score 133; DB 4; Length 484;
 Best Local Similarity 26.2%; Pred. No. 4.3e-05; Mismatches 81; Indels 32; Gaps 9;
 Matches 53; Conservative 36; Gaps 9;

Qy 408 MRALYIVRDPRAWIYSMLYNKSPLSY--SLKNVPE----HLAKLFLEGK-GKCNLNS 459
 Db 279 LKVTHLVRDRA-VASSRISRHGLIRESQVRSRDPRAHRMPFLEAIGHLGAKKEGV 337

Query Match 4.2%; Score 133; DB 4; Length 484;
 Best Local Similarity 26.2%; Pred. No. 4.3e-05; Mismatches 81; Indels 32; Gaps 9;
 Matches 53; Conservative 36; Gaps 9;

Qy 460 GYAFYEPLRKELSKSKSNAVSLSHLWANTAAALRINTDLPLTSYOLVKFEDIVHFPQ 519
 Db 338 GGPDYHAL-----GAMEVICNSMAKTLQTAQ-PDDWLOGHLYLVRVEDLVGDPV 387

Query Match 4.2%; Score 133; DB 4; Length 484;
 Best Local Similarity 26.2%; Pred. No. 4.3e-05; Mismatches 81; Indels 32; Gaps 9;
 Matches 53; Conservative 36; Gaps 9;

Qy 520 KTTERIFAGIPLSPA---SINOILPATSTNLFLYLYGEISPTN----TNWKONLP 571
 Db 388 KTLRVDYDFGLVSPMEQFALNMTCGSSSKPFPV----VSARNATOAAANAWRTALT 442

Query Match 4.2%; Score 133; DB 4; Length 484;
 Best Local Similarity 26.2%; Pred. No. 4.3e-05; Mismatches 81; Indels 32; Gaps 9;
 Matches 53; Conservative 36; Gaps 9;

Qy 572 RDEIKLIENICWTMLDRGPK 593
 Db 443 FQQIKQVEEFCYQPMNAVLYER 464

RESULT 7
 US-09-949-016-6471
 ; Sequence 6471, Application US/09949016
 ; Patent No. 6812339

; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 6471
 ; LENGTH: 531
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-949-016-6471

Query Match 4.2%; Score 133; DB 4; Length 531;
 Best Local Similarity 26.2%; Pred. No. 5e-05; Mismatches 81; Indels 32; Gaps 9;
 Matches 53; Conservative 36; Gaps 9;

Qy 408 MRALYIVRDPRAWIYSMLYNKSPLSY--SLKNVPE----HLAKLFLEGK-GKCNLNS 459
 Db 326 LKVTHLVRDRA-VASRIRSRHGLIRESQVRSRDPRAHRMPFLEAIGHLGAKKEGV 384

RESULT 6
 US-09-471-867-4
 ; Sequence 4, Application US/09471867
 ; Patent No. 6455289
 ; GENERAL INFORMATION:
 ; APPLICANT: Uchimura, Kenji
 ; APPLICANT: Muramatsu, Hideki
 ; APPLICANT: Kannagi, Reiji
 ; APPLICANT: Habuchi, Osami

RESULT 8

Qy 460 GYAFYEPLRKELSKSKSNAVSLIHLWLTANTAAALRINTDILPPTSYOLVKPEDIYHFPQ 519 ; PRIORITY APPLICATION NUMBER: 60/237,768
 Db 385 GGPDYHAL-----GAEVICNSMAKTLQALQ-PPDWLQGHYLIVRYEDLVGDPV 434 ; PRIORITY FILING DATE: 2000-10-03
 Qy 520 KITERIFAGIPISPA---SINOILPATSTNLFLPYEGEISPTN----TNUWKQNLP 571 ; PRIORITY APPLICATION NUMBER: 60/231,498
 Db 435 KTLLRVDYDFGLVSPMEEQFALNMTSGGGSSSKPFV----VSARNTQAANAWRTALT 489 ; NUMBER OF SEQ ID NOS: 207012
 Qy 572 RDEIKLIENTICWTLMDRGYPK 593 ; SOFTWARE: FastSEQ For Windows Version 4.0
 Db 490 FQQIKQVEEFYCQPMAVLYER 511 ; SEQ ID NO: 7011
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-7011

Query Match Score 3.98; Length 395;
 Best Local Similarity 23.3%; Pred. No. 0.00012;
 Matches 90; Conservative 53; Mismatches 140; Indels 103; Gaps 20;

Qy 239 VVITSLPGSGABILKQFLFNSSDFLYTRVPAYIDFETELEBDSYACENKV-SPIRS 297
 Db 43 VLVLSSRSRGSSSFVGQFNQHPDVFTIMEPA-----WNVWTTLSQ 82

Qy 298 GHFRLLJFGWLIQSLVQDUTKL-HIQNITHLHEPNRNGKL-AQYFAMNKDKKKFKRRESLPEORS 356
 Db 83 GSRATLIMAARDLVRSFLCDMDVFLDYLPRNRNLSSLF-----QWAVSRALCSPAC 136

Qy 357 QMKGAFDRAFTDAEYTRALRHLVYPPS--ARPVLSLSSGSWTLKLHFFQ-EVL----G 405
 Db 137 ----AFRGAISSEAKPLCAROFSTLAREACRSHVYKEVRPNLQVLYPLSDPA 192

Qy 406 ASMRALYIVRDPAWYIMSYLNKSPLSYLSKVNPEHIAKLFIEGG--KGKCNLNSSYAF 463
 Db 193 LNRLRIVLVRDPAVLRSR-----EQATAKALARNDIVLG----TNGTW 233

Qy 464 EYEP-LR--KELSKSKSNAVSLIHLWLTANTAAALRINTDILPPTSYOLVKPEDIYHFPQ 520
 Db 234 EADPGLVRREYCR-----SHVRIAE-AATLK-PPPFGRGTYLVRFDLARPSLA 282

Qy 521 TTERIFAGIPLSP--ASLNOI-----LATATSNLFLYEGEISPTN-TNV 565
 Db 283 EIRALYFTGSLUTPOLEAWHTNIGSGPGRAREAKTSSR-----NALNSYQA 332

RESULT 10

US-08-699-514-2

Qy 566 WKNLPRDEIKLJENICWTLMDRGY 591 ; Sequence 2, Application US/06899514
 Db 333 WRHALPAKIRRVQELCAGALQGY 358 ; Patent No. 5910581

GENERAL INFORMATION:

APPLICANT: HABUCHI, OSAMU
 ADDRESS: KOBBE, MARPENS, OLSON & BEAR, LLP
 STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
 CITY: NEWPORT BEACH
 STATE: CALIFORNIA
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/699,514
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

RESULT 13
 US-09-786-240-11
 ; Sequence 11; Application US/09786240
 ; Patent No. 6556935
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE PHARMACEUTICALS, INC.
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: CORLEY, Neil C.
 ; APPLICANT: DOUGLER, Karl J.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: LALI, Preeti
 ; APPLICANT: YUE, Harry
 ; APPLICANT: HILLMAN, Jennifer L.
 ; APPLICANT: AZIMZAI, Yalda
 ; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 ; FILE REFERENCE: PF-0592_PCT
 ; CURRENT APPLICATION NUMBER: US/09786_240
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
 ; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1999-05-11
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PERL Program
 ; SEQ ID NO: 11
 ; LENGTH: 386
 ; OTHER INFORMATION: Incyte ID No. 6556935 2617407CD1
 ; US-09-786-240-11

Query Match Score 3.7%; Score 115.5%; DB 4; Length 386;
 Best Local Similarity 20.4%; Prod. No. 0.0023; Mismatches 151; Indels 117; Gaps 17;
 Matches 83; Conservative 56; Name/Key: misc feature

Db 129 TLHANVARDLMSRISFL--CDMDVFDAYMQSRNLNSAFFNWATSA-----LCSPP 175
 Qy 327 ---NRGKLAQYFAMNKDKRKFKRRESLPEQRSMQKGAFDRDAEYTRARRHLVYPS 381
 Db 176 ACSAFPRTISK----QDVCKTLLCITQPFSFLAREACRSYSHVLKHFVFLQVLYPLL 230
 Qy 382 ARPVLSLSSGSWTLKLHPPQEVLGASHMRALYVRDPWAIYKMLYNSSKPSLSSLKVNPEH 441
 Db 231 SDPALNL-----RIVLVYDPAVLRER-EAAGPILARDNGI--- 266
 Qy 442 LAKLFKTEGGKGKCNLNSSGYAPEYEP---LRLKELSKSKSNASVSLSSHMLWANTAAIRIN 498
 Db 267 -----VLGTNGRK-----VEADPHLRLIREVCV-----SHVRAE-AATLK-P 302

RESULT 14
 US-09-045-284A-2
 ; Sequence 2; Application US/09045284A
 ; Patent No. 6265192
 ; GENERAL INFORMATION:
 ; APPLICANT: Bistrup, Annette
 ; APPLICANT: Rosen, Steven D.
 ; APPLICANT: Hemmerich, Stefan
 ; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
 ; FILE REFERENCE: 6510-10707S1
 ; CURRENT APPLICATION NUMBER: US/09/045, 284A
 ; CURRENT FILING DATE: 1998-03-20
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 2
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-045-284A-2

Query Match Score 3.4%; Score 108.5%; DB 3; Length 386;
 Best Local Similarity 18.9%; Prod. No. 0.013; Mismatches 55; Indels 129; Gaps 15;
 Matches 78; Conservative 55; Name/Key: Glycosyl Sulftotransferase-3

Db 303 PPFLRGYRVLRFEDLARBLPBLAFLRUYAFTGLTLPQLEAWIHNTHGSGIGKPTEAFH 362
 Qy 558 ISPTN---INTWKQNLPDRBIKLIENICWTLMRDGY 591
 Db 363 TSSRNARNVSQAMRHAPFTKLRLVQEVCAGALQLGY 400

RESULT 15
 US-10-190-911-1
 ; Sequence 1; Application US/09190911
 ; Patent No. 6365365
 ; GENERAL INFORMATION:

```

; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Tandemann, Kirsten
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE - 3
; FILE REFERENCE: 6510-107CIP
; CURRENT APPLICATION NUMBER: US/09/190,911
; CURRENT FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09,045,284
; EARLIER FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 386
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-190-911-1

Query Match          3 4%; Score 108 5; DB 3; Length 386;
Best Local Similarity 18.9%; Pred. No. 0.013; Gaps 15;
Matches 78; Conservative 55; Mismatches 151; Indels 129; Gaps 15;
Qy   223 SKKSUSSEGHHMDLPDVITSLPSSGAEIJKQLFNSDPELYIRVPTAYTIDIPETELEID 282
Db   32 SLSMKAAQPERM---HVLVLSWSSGSSTFGQLFCQHPDWYLMEPAMHV-----78
Qy   283 SFVDACEWKYSDIRSGHFRLLRGWILQSLVDTKLHLQNLTHHEPNRGLKLAQYFAMNKDKK 342
Db   79 -----WMT-----FKCTAWN-----LHMAYRDLIRAVFLCDMSVF 109
Qy   343 RKEFKRRESLPEORSMQKGAFDRDAEYIRAL-----RRHIVYYPSARPVIL 386
Db   110 DAY--MEPGPRROSL---FOWENSRALCSACADITPQDEIIIPRAHCRLCSQQPFE 162
Qy   387 SLSSSSWTLLKHFPEVILGASMRLY-----I.VRDPRAWIYMSLYNSKPSL 432
Db   163 VVEKACRSYSHVVLKEVRFRFLQSLSYPLKDKPSLNLIHIVLVRDRA---VRSR--- 214
Qy   433 YSLKNVPPEHIAKLFIEGGKGCKNUNLNSGYAF--EYBPLRKELSKSKSNAVSLISHLWLN 490
Db   215 -----ERTGDLMDSDR1VNGQHHRQKLXKBEDQPYVVMOTICQSOLEYIK 258
Qy   491 TAAALRINTDILPLTSYQLVKPEDITHPQKTTTERIPLAFLGIPLSPASLNQILEFT---- 545
Db   259 TIQSL---PKALQERYLLVRYEDILARAPVQTSRMYEFVYLEFLPHLOTWHNTRGKRM 315
Qy   546 STNLFPYLPYGEISPTN-----TNWIKONLPRDEIKLBNICWTLMDLGY 591
Qy   316 GDHATH-----TNARDALNVSOAWRNSLPYEKVSRLOKACGDMNLGY 359

```

THIS PAGE BLANK (USPTO)

| Result | No. | Score | Query | Match | Length | DB | ID | Description | Sequence |
|--------|------|-------|-------|------------------|--------------------|--------------------|----|-------------|--------------------|
| - | - | - | - | - | - | - | - | - | - |
| 1 | 3147 | 100.0 | 596 | 16 | US-10-697-828-9 | Sequence 9, Appl1 | | | Sequence 1544, AP |
| 2 | 3147 | 100.0 | 1222 | 16 | US-10-697-828-15 | Sequence 15, Appl1 | | | Sequence 2, Appl1 |
| 3 | 3142 | 99.8 | 1212 | 17 | US-10-697-828-15 | Sequence 2, Appl1 | | | Sequence 4, Appl1 |
| 4 | 3142 | 99.8 | 1222 | 16 | US-10-697-828-15 | Sequence 4, Appl1 | | | Sequence 28*, App |
| 5 | 2835 | 90.1 | 1207 | 16 | US-10-697-828-17 | Sequence 17, Appl1 | | | Sequence 2, Appl1 |
| 6 | 2835 | 90.1 | 1207 | 17 | US-10-697-828-17 | Sequence 4, Appl1 | | | Sequence 5, Appl1 |
| 7 | 568 | 125 | 11 | US-09-984-429-71 | Sequence 71, Appl1 | | | | Sequence 159, App |
| 8 | 568 | 18.0 | 126 | 14 | US-10-150-111-71 | Sequence 71, Appl1 | | | Sequence 128, App |
| 9 | 147 | 4.7 | 411 | 14 | US-10-021-462-97 | Sequence 97, Appl1 | | | Sequence 97, Appl1 |
| 10 | 147 | 4.7 | 411 | 15 | US-10-021-462-97 | Sequence 395, App | | | Sequence 395, App |
| 11 | 147 | 4.7 | 411 | 16 | US-10-408-765A-395 | | | | |

| | | |
|---|---|---|
| Copyright (c) 1993 - 2005 Compugen Ltd. | GenCore version 5.1.6 | Sequence 1544, AP |
| OM protein - protein search, using SW model | Sequence 2, Appl1 | Sequence 2, Appl1 |
| Run on: June 23, 2005, 08:49:14 ; Search time 75.3449 Seconds (without alignments) | Sequence 4, Appl1 | Sequence 4, Appl1 |
| Title: US-10-697-828-9 | Sequence 255, App | Sequence 255, App |
| Perfect score: 3147 | Sequence 28*, App | Sequence 28*, App |
| Sequence: 1 MNRYNGAMMDVWDAHYKMFNWL.....LIENICWTLMDRGYPKFMD 596 | Sequence 2, Appl1 | Sequence 2, Appl1 |
| Scoring table: BIOSUM62 | Sequence 5, Appl1 | Sequence 5, Appl1 |
| Gapop 10.0 , Gapext 0.5 | Sequence 126, App | Sequence 126, App |
| Searched: 1717557 seqs, 384547976 residues | Sequence 4, Appl1 | Sequence 4, Appl1 |
| Total number of hits satisfying chosen parameters: 1717557 | Sequence 8, Appl1 | Sequence 8, Appl1 |
| Minimum DB seq length: 0 | Sequence 11, Appl1 | Sequence 11, Appl1 |
| Maximum DB seq length: 2000000000 | Sequence 3, Appl1 | Sequence 3, Appl1 |
| Post-processing: Minimum Match 0% Maximum Match 10% | Sequence 6, Appl1 | Sequence 6, Appl1 |
| Listing first 45 summaries | Sequence 824, App | Sequence 824, App |
| Database : Published Applications AA:* | Sequence 2, Appl1 | Sequence 2, Appl1 |
| 1: /cg92_6/podata/1/pubpa/PUBCOMB.pep:* | Sequence 1, Appl1 | Sequence 1, Appl1 |
| 2: /cg92_6/podata/1/pubpa/PCT_NEW_PUB..pep:* | Sequence 17, Appl1 | Sequence 17, Appl1 |
| 3: /cg92_6/podata/1/pubpa/US06_PUBCOMB.pep:* | Sequence 7118, Ap | Sequence 7118, Ap |
| 4: /cg92_6/podata/1/pubpa/US05_PUBCOMB.pep:* | Sequence 73555, A | Sequence 73555, A |
| 5: /cg92_6/podata/1/pubpa/US07_PUBCOMB.pep:* | Sequence 104463, | Sequence 104463, |
| 6: /cg92_6/podata/1/pubpa/US08_PUBCOMB.pep:* | Sequence 9, Application US/10697828-9 | Sequence 9, Application US/10697828-9 |
| 7: /cg92_6/podata/1/pubpa/US09_PUBCOMB.pep:* | Publication No. US20040185546A1 | Publication No. US20040185546A1 |
| 8: /cg92_6/podata/1/pubpa/US09B_PUBCOMB.pep:* | GENERAL INFORMATION: | GENERAL INFORMATION: |
| 9: /cg92_6/podata/1/pubpa/US09C_PUBCOMB.pep:* | ; APPLICANT: Rosen, Steven | ; APPLICANT: Rosen, Steven |
| 10: /cg92_6/podata/1/pubpa/US09D_PUBCOMB.pep:* | ; ATTORNEY: Lee, Jin Kyu | ; ATTORNEY: Lee, Jin Kyu |
| 11: /cg92_6/podata/1/pubpa/US09E_PUBCOMB.pep:* | ; HAMMICH, Stefan | ; HAMMICH, Stefan |
| 12: /cg92_6/podata/1/pubpa/US09F_PUBCOMB.pep:* | ; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-6 | ; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-6 |
| 13: /cg92_6/podata/1/pubpa/US10A_PUBCOMB.pep:* | FILE REFERENCE: UCAL-138D1V | FILE REFERENCE: UCAL-138D1V |
| 14: /cg92_6/podata/1/pubpa/US10B_PUBCOMB.pep:* | CURRENT APPLICATION NUMBER: US/10-697, 828 | CURRENT APPLICATION NUMBER: US/10-697, 828 |
| 15: /cg92_6/podata/1/pubpa/US10C_PUBCOMB.pep:* | PRIOR APPLICATION NUMBER: 09-593, 828 | PRIOR APPLICATION NUMBER: 09-593, 828 |
| 16: /cg92_6/podata/1/pubpa/US10D_PUBCOMB.pep:* | PRIOR FILING DATE: 2000-06-13 | PRIOR FILING DATE: 2000-06-13 |
| 17: /cg92_6/podata/1/pubpa/US10E_PUBCOMB.pep:* | PRIOR APPLICATION NUMBER: 60-144, 694 | PRIOR APPLICATION NUMBER: 60-144, 694 |
| 18: /cg92_6/podata/1/pubpa/US10F_PUBCOMB.pep:* | PRIOR FILING DATE: 1999-07-20 | PRIOR FILING DATE: 1999-07-20 |
| 19: /cg92_6/podata/1/pubpa/US11A_PUBCOMB.pep:* | NUMBER OF SEQ ID NOS: 23 | NUMBER OF SEQ ID NOS: 23 |
| 20: /cg92_6/podata/1/pubpa/US11B_PUBCOMB.pep:* | SOFTWARE: FastSQL for Windows Version 4.0 | SOFTWARE: FastSQL for Windows Version 4.0 |
| 21: /cg92_6/podata/1/pubpa/US11C_PUBCOMB.pep:* | SEQ ID NO: 9 | SEQ ID NO: 9 |
| 22: /cg92_6/podata/1/pubpa/US11D_PUBCOMB.pep:* | LENGTH: 596 | LENGTH: 596 |
| | TYPE: PRT | TYPE: PRT |
| | ORGANISM: homo sapiens | ORGANISM: homo sapiens |

RESULT 1
US-10-697-828-9
; Sequence 9, Application US/10697828-9
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven
; ATTORNEY: Lee, Jin Kyu
; HAMMICH, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-6
; FILE REFERENCE: UCAL-138D1V
; CURRENT FILING DATE: 2003-10-29
; PRIORITY FILING DATE: 2000-06-13
; PRIORITY APPLICATION NUMBER: 09-593, 828
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO: 9
; LENGTH: 596

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query | Match | Length | DB | ID | Description | Sequence |
|--------|------|-------|-------|------------------|--------------------|--------------------|----|-------------|--------------------|
| - | - | - | - | - | - | - | - | - | - |
| 1 | 3147 | 100.0 | 596 | 16 | US-10-697-828-9 | Sequence 9, Appl1 | | | Sequence 1544, AP |
| 2 | 3147 | 100.0 | 1222 | 16 | US-10-697-828-15 | Sequence 15, Appl1 | | | Sequence 2, Appl1 |
| 3 | 3142 | 99.8 | 1212 | 17 | US-10-697-828-15 | Sequence 2, Appl1 | | | Sequence 4, Appl1 |
| 4 | 3142 | 99.8 | 1222 | 16 | US-10-697-828-15 | Sequence 4, Appl1 | | | Sequence 28*, App |
| 5 | 2835 | 90.1 | 1207 | 16 | US-10-697-828-17 | Sequence 17, Appl1 | | | Sequence 2, Appl1 |
| 6 | 2835 | 90.1 | 1207 | 17 | US-10-697-828-17 | Sequence 4, Appl1 | | | Sequence 5, Appl1 |
| 7 | 568 | 125 | 11 | US-09-984-429-71 | Sequence 71, Appl1 | | | | Sequence 159, App |
| 8 | 568 | 18.0 | 126 | 14 | US-10-150-111-71 | Sequence 71, Appl1 | | | Sequence 128, App |
| 9 | 147 | 4.7 | 411 | 14 | US-10-021-462-97 | Sequence 97, Appl1 | | | Sequence 97, Appl1 |
| 10 | 147 | 4.7 | 411 | 15 | US-10-021-462-97 | Sequence 395, App | | | Sequence 395, App |
| 11 | 147 | 4.7 | 411 | 16 | US-10-408-765A-395 | | | | |

Query Match 100.0%; Best Local Similarity 100.0%; Matches 596; Conservativeness 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARYNGAMMDWDHYKMFWDHGNSPMASIQAEQAEFRKKEWTOYFNVTFOEMESTIT 60

Db 1 MARYNGAMMDWDHYKMFWDHGNSPMASIQAEQAEFRKKEWTOYFNVTFOEMESTIT 60

Qy 61 RIAYVFGPYTINVSCRFIDSSNPGQLQISIVNNTEVHSIVTVDYHNLKTRNLYLGFGF 120

| | | | |
|----|--|----|---|
| Db | 61 RIAVVFYGPVINYSSCRFIDSSNPGQISSLVNNTHEHVSIVTDYHNLKTRFNYLGGFF 1.20 | Db | 747 ASVADQGQITRFGTQAIKVPKVRHDIIIPPGFKENIAVGVLICSLVLTQWRRFLS 806 |
| Qy | 121 ASVADQGQITRFGTQAIKVPKVRHDIIIPPGFKENIAVGVLICSLVLTQWRRFLS 1.80 | Qy | 181 FRKLMPRLILYLTAIWPLFELDQISSLVNTHEHVSIVTDYHNLKTRFNYLGGFF 240 |
| Db | 121 ASVADQGQITRFGTQAIKVPKVRHDIIIPPGFKENIAVGVLICSLVLTQWRRFLS 1.80 | Db | 807 FRKLMPRLILYLTAIWPLFELDQISSLVNTHEHVSIVTDYHNLKTRFNYLGGFF 866 |
| Qy | 181 FRKLMPRLILYLTAIWPLFELDQISSLVNTHEHVSIVTDYHNLKTRFNYLGGFF 240 | Qy | 241 ITSLPGSGBAIKIKOLFNSSDLYIRPTAYIDIPETELEIDSFDACEWKYSDIRCHF 3.00 |
| Db | 181 FRKLMPRLILYLTAIWPLFELDQISSLVNTHEHVSIVTDYHNLKTRFNYLGGFF 240 | Db | 867 ITSLPGSGBAIKIKOLFNSSDLYIRPTAYIDIPETELEIDSFDACEWKYSDIRCHF 926 |
| Qy | 241 ITSLPGSGBAIKIKOLFNSSDLYIRPTAYIDIPETELEIDSFDACEWKYSDIRCHF 3.00 | Qy | 301 RLIRGMWLSYDQDTKLHLQNTLHLPEPRGKLAQYFAMNKDKCRKFKRESLPEQRSMKG 3.60 |
| Db | 241 ITSLPGSGBAIKIKOLFNSSDLYIRPTAYIDIPETELEIDSFDACEWKYSDIRCHF 3.00 | Db | 927 RLIRGMWLSYDQDTKLHLQNTLHLPEPRGKLAQYFAMNKDKCRKFKRESLPEQRSMKG 986 |
| Qy | 301 RLIRGMWLSYDQDTKLHLQNTLHLPEPRGKLAQYFAMNKDKCRKFKRESLPEQRSMKG 3.60 | Qy | 361 AFRDADYIRALPRLVWYPSARPVLSLSSGSMBALYVDRDRAW 420 |
| Db | 301 RLIRGMWLSYDQDTKLHLQNTLHLPEPRGKLAQYFAMNKDKCRKFKRESLPEQRSMKG 3.60 | Db | 987 AFRDADYIRALRRLVWYPSARPVLSLSSGSMRALYVDRDRAW 1046 |
| Qy | 361 AFRDADYIRALPRLVWYPSARPVLSLSSGSMBALYVDRDRAW 420 | Qy | 421 IYSLMLYNKSPSLSYLKNPVEHLAKLFLIEGGKCNLNSGFAEYPLRKELSKSKNAV 4.80 |
| Db | 361 AFRDADYIRALPRLVWYPSARPVLSLSSGSMBALYVDRDRAW 420 | Db | 1047 IYSLMLYNKSPSLSYLKNPVEHLAKLFLIEGGKCNLNSGFAEYPLRKELSKSKNAV 1106 |
| Qy | 421 IYSLMLYNKSPSLSYLKNPVEHLAKLFLIEGGKCNLNSGFAEYPLRKELSKSKNAV 4.80 | Qy | 481 SLSLHLWLNATAALRINTDLIPLTSYOLVKFEDIVHPHQKTERIFAPLGLIPSPASLNO 5.40 |
| Db | 421 IYSLMLYNKSPSLSYLKNPVEHLAKLFLIEGGKCNLNSGFAEYPLRKELSKSKNAV 4.80 | Db | 1107 SLSLHLWLNATAALRINTDLIPLTSYOLVKFEDIVHPHQKTERIFAPLGLIPSPASLNO 1166 |
| Qy | 481 SLSLHLWLNATAALRINTDLIPLTSYOLVKFEDIVHPHQKTERIFAPLGLIPSPASLNO 5.40 | Qy | 541 ILFATSTNLFLYLPYEGISPTNTNVKONLPRDEIJKLJENICWTLMRGLYPKEMD 5.96 |
| Db | 481 SLSLHLWLNATAALRINTDLIPLTSYOLVKFEDIVHPHQKTERIFAPLGLIPSPASLNO 5.40 | Db | 1167 ILFATSTNLFLYLPYEGISPTNTNVKONLPRDEIJKLJENICWTLMRGLYPKEMD 1222 |
| Qy | 541 ILFATSTNLFLYLPYEGISPTNTNVKONLPRDEIJKLJENICWTLMRGLYPKEMD 5.96 | | RESULT 3 |
| Db | 541 ILFATSTNLFLYLPYEGISPTNTNVKONLPRDEIJKLJENICWTLMRGLYPKEMD 5.96 | | US-10-479-472A-2 |
| | | | ; Sequence 2, Application US/10479472A |
| | | | ; GENERAL INFORMATION: |
| | | | ; PUBLICATION NO. US2005011858A1 |
| | | | ; APPLICANT: DEL-FAVERO, JURGEN PETER LODE |
| | | | ; ADDRESS: VAN BROEKHOVEN CHRISTINE |
| | | | ; TITLE OF INVENTION: NOVEL BRAIN EXPRESSED GENE AND PROTEIN ASSOCIATED WITH BIPOLAR DISORDER |
| | | | ; FILE REFERENCE: JAB-1711 |
| | | | ; CURRENT APPLICATION NUMBER: US/10/479,472A |
| | | | ; CURRENT FILING DATE: 2003-12-01 |
| | | | ; PRIOR APPLICATION NUMBER: PCT/EP02/06316 |
| | | | ; PRIOR FILING DATE: 2002-06-06 |
| | | | ; PRIOR APPLICATION NUMBER: EP 01202214.1 |
| | | | ; PRIOR FILING DATE: 2001-06-11 |
| | | | ; NUMBER OF SEQ ID NOS: 12 |
| | | | ; SOFTWARE: Patentin Ver. 3.2 |
| | | | ; SEQ ID NO 2 |
| | | | ; LENGTH: 1212 |
| | | | ; TYPE: PRT |
| | | | ; ORGANISM: Homo sapiens |
| | | | US-10-479-472A-2 |
| | | | RESULT 2 |
| | | | US-10-697-828-15 |
| | | | ; Sequence 15, Application US/10697828 |
| | | | ; GENERAL INFORMATION: |
| | | | ; PUBLICATION NO. US20040185546A1 |
| | | | ; APPLICANT: Hemmerich, Stefan |
| | | | ; ADDRESS: Lee, Jin Kyu |
| | | | ; TITLE OF INVENTION: Novel Glycosyl Sulfolotransferases GST-4alpha, GST-4beta, & GST-6 |
| | | | ; FILE REFERENCE: UCAU-138BIV |
| | | | ; CURRENT APPLICATION NUMBER: US/10/697,828 |
| | | | ; CURRENT FILING DATE: 2003-10-29 |
| | | | ; PRIOR APPLICATION NUMBER: 09/593,828 |
| | | | ; PRIOR FILING DATE: 2000-06-13 |
| | | | ; PRIOR APPLICATION NUMBER: 60/144,694 |
| | | | ; PRIOR FILING DATE: 1999-07-20 |
| | | | ; NUMBER OF SEQ ID NOS: 23 |
| | | | ; SOFTWARE: FastSEQ for Windows Version 4.0 |
| | | | ; SEQ ID NO 15. |
| | | | ; LENGTH: 1.22 |
| | | | ; TYPE: PRT |
| | | | ; ORGANISM: homo sapiens |
| | | | US-10-697-828-15 |
| | | | Query Match 100.0%; Score 3147; DB 16; Length 1222; |
| | | | Best Local Similarity 100.0%; Pred. No. 7.6e-295; Mismatches 0; Indels 0; Gaps 0; |
| | | | Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| Qy | 1 MNRYNGAMMDWDHYKPFWDHGNSPMASIQAEQAFKRWTOVNNTFQMESTIT 60 | Qy | 1 MNRYNGAMMDWDHYKPFWDHGNSPMASIQAEQAFKRWTOVNNTFQMESTIT 60 |
| Db | 627 MNRYNGAMMDWDHYKPFWDHGNSPMASIQAEQAFKRWTOVNNTFQMESTIT 686 | Db | 617 MNRYNGAMMDWDHYKPFWDHGNSPMASIQAEQAFKRWTOVNNTFQMESTIT 676 |
| Qy | 61 RIAVVFYGPVINYSSCRFIDSSNPGQISSLVNNTHEHVSIVTDYHNLKTRFNYLGGFF 120 | Qy | 61 RIAVVFYGPVINYSSCRFIDSSNPGQISSLVNNTHEHVSIVTDYHNLKTRFNYLGGFF 120 |
| Db | 687 RIAVVFYGPVINYSSCRFIDSSNPGQISSLVNNTHEHVSIVTDYHNLKTRFNYLGGFF 746 | Db | 677 RIAVVFYGPVINYSSCRFIDSSNPGQISSLVNNTHEHVSIVTDYHNLKTRFNYLGGFF 736 |
| Qy | 121 ASVADQGQITRFGTQAIKVPKVRHDIIIPPGFKENIAVGVLICSLVLTQWRRFLS 180 | Qy | 121 ASVADQGQITRFGTQAIKVPKVRHDIIIPPGFKENIAVGVLICSLVLTQWRRFLS 180 |
| Db | | Db | 737 ASVADQGQITRFGTQAIKVPKVRHDIIIPPGFKENIAVGVLICSLVLTQWRRFLS 796 |
| Qy | | Qy | 181 FRKLMPRLILYLTAIWPLFELDQISSLVNTHEHVSIVTDYHNLKTRFNYLGGFF 240 |

RESULT 4
US-10-475-446-4
; Sequence 4, Application US/10475446
; Publication No. US20050198551A1
; GENERAL INFORMATION:
; APPLICANT: Klammer, Aaron A.; Hafajta, April J.A.
; DUGGAN, Brendan M.; WARREN, Bridget A.
; APPLICANT: ARVIZU, Chandra S.; HONCHELL, Catherine M.
; NGUYEN, Danniell B.; KALLICK, Deborah A.
; APPLICANT: YUE, Henry; Janice K.
; RAMKUMAR, Jayalakshmi; Li, Joana X.
; APPLICANT: THANGAVELU, Ravitha; Gietzen, Kimberly J.
; DING, Li; BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.; CHAWLA, Narinder K.
; APPLICANT: MASON, Patricia M.; LAL, Preeti G.
; APPLICANT: GRAUL, Richard P.; KAREHT, Stephanie K.
; BECHA, Shanya D.; RICHARDSON, Thomas W.; TRAN, Uyen K.
; APPLICANT: ELLIOTT, Vicki S.; TANG, Y.; Tom
; APPLICANT: AZIMZAI, Yalda; LU, Yan
; APPLICANT: XU, Yuming
; TITLE OF INVENTION: SECRETED PROTEINS
; FILE NUMBER: PF-0349 USN
; CURRENT APPLICATION NUMBER: US 10/475,446
; PRIOR APPLICATION NUMBER: PCT/US02/12464
; PRIOR FILING DATE: 2003-10-20
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: US 60/285,207
; PRIOR FILING DATE: 2001-04-20
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/288,640
; PRIOR FILING DATE: 2001-05-03
; PRIOR FILING DATE: 2001-05-11
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/343,553
; PRIOR FILING DATE: 2001-12-21
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/358,279
; PRIOR FILING DATE: 2002-03-19

RESULT 5
US-10-697-828-17
; Sequence 17, Application US/10697828
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Roser, Steven
; APPLICANT: Lee, Jin Kyu
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-6
; FILE REFERENCE: UCAL-138D1V
; CURRENT APPLICATION NUMBER: US 10/697,828
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: 09 593,828
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/144,694
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS : 23

Db 797 FRKLMRWLILVIAWFIELDWSTCSQPICAKWTRTEAGSKKSLSSEGHMHMDLPDV 856 ; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PERL Program
; SEQ ID NO: 4
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 3441255CD1
; US-10-475-446-4
Query Match 99.8%; Score 3142; DB 16; Length 1222;
Best Local Similarity 99.8%; Pred. No. 2_3e-294;
Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MNYTNGAMMDWDAHYKNEWFDHGNSPMASIGEAEQAAEFKKRWTQFVNFTFOMESETIT 60
Db 627 MNRYTNGAMMDWDAHYKNEWFDHGNSPMASIGEAEQAAEFKKRWTQFVNFTFOMEPTIT 686
Qy 61 RIAVYFGPYINNSCRFIDSSNPGLQISLNVNNTENHVSIVTDYHNLKTRENVLGFGGP 120
Db 687 RIAVYFGPYINNSCRFIDSSNPGLQISLNVNNTENHVSIVTDYHNLKTRENVLGFGGP 746
Qy 121 ASVADQGQTTRFLGTQATVKPVYRDR1IFPFGKFNIAVGLLICI SVLTLTFLWRFYL 180
Db 747 ASVADQGQTTRFLGTQATVKPVYRDR1IFPFGKFNIAVGLLICI SVLTLTFLWRFYL 806
Qy 181 FRKLMRWLILVIALWFELLDWSTSCOPICAKWTRTEAGSCKSLSSSEGHNLDLPVV 240
Db 807 FRKLMRWLILVIALWFELLDWSTSCOPICAKWTRTEAGSCKSLSSSEGHNLDLPVV 866
Qy 241 ITSLPGSGAEBILKOLFNSDFLYIRVPPAYIDIPETELEIDSVDACEMKVSDIRSGHF 300
Db 867 ITSLPGSGAEBILKOLFNSDFLYIRVPPAYIDIPETELEIDSVDACEMKVSDIRSGHF 926
Qy 301 RLLRGWLQSLVQDVTKLHLQNTLHBPNRGKLAQYFAMNKDKRKFRKRSLSPEORS QMKG 360
Db 927 RLLRGWLQSLVQDVTKLHLQNTLHBPNRGKLAQYFAMNKDKRKFRKRSLSPEORS QMKG 986
Qy 361 AFDDAEYTRALRHLYVPSAREVLSSLSGGSMWTLKLIFFQEVGSMRALYTRDPRW 420
Db 987 AFDDAEYTRALRHLYVPSAREVLSSLSGGSMWTLKLIFFQEVGSMRALYTRDPRW 1046
Qy 421 IYSMLYNKSPSLYSLKVNPFELHLKFKEGGKGKCNLNGYAFPEPLRKELSKSKSNAV 480
Db 1047 IYSMLYNKSPSLYSLKVNPFELHLKFKEGGKGKCNLNGYAFPEPLRKELSKSKSNAV 1106
Qy 481 SLLSHLWLANTAAALRINTDLPPTSYQLVKFDITVHPQTTERIAFGLPGLPSASLNQ 540
Db 1107 SLLSHLWLANTAAALRINTDLPPTSYQLVKFDITVHPQTTERIAFGLPGLPSASLNQ 1166

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 17

; LENGTH: 1207

; TYPE: PRT

; ORGANISM: mouse

US-10-697-828-17

; ORGANISM: Mus sp.

US-10-479-472A-4

Query Match 90.1%; Score 2835; DB 16; Length 1207;

Best Local Similarity 89.6%; Pred. No. 1.3e-264;

Mismatches 36; Indels 0; Gaps 0;

Matches 528; Conservative 36; Mismatches 32; Indels 0; Gaps 0;

Query 1 MNRYNGAMMDWDHYKMWFDHHGNSPMASIQAEQAAEKKRWTQFVNNTFQMESTIT 60

Db 612 MNRYNGAMMDWDHYKMWFDHHGNSPMASIQAEQAAEKKRWTQFVNNTFQMESTIT 60

Query 61 RIAVYFGPYVINSSCRFIDSSNGLQISLNNTTEHVSIYTDYHNLKTRENYKTFVNNTFQMESTIT 60

Db 671 RIAVYFGPYVINSSCRFIDSSNGLQISLNNTTEHVSIYTDYHNLKTRENYKTFVNNTFQMESTIT 671

Query 61 RIAVYFGPYVINSSCRFIDSSNGLQISLNNTTEHVSIYTDYHNLKTRENYKTFVNNTFQMESTIT 60

Db 672 RIAVYFGPYVINSSCRFIDSSNGLQISLNNTTEHVSIYTDYHNLKTRENYKTFVNNTFQMESTIT 671

Query 61 RIAVYFGPYVINSSCRFIDSSNGLQISLNNTTEHVSIYTDYHNLKTRENYKTFVNNTFQMESTIT 60

Db 672 RIAVYFGPYVINSSCRFIDSSNGLQISLNNTTEHVSIYTDYHNLKTRENYKTFVNNTFQMESTIT 671

Query 121 ASVANQGQITRFLGTOAIKVPRHRDRLLPFGFKENIAGVGLICSLIVLTQWRFLS 180

Db 732 ASVANQGQITRFLGTOAIKVPRHRDRLLPFGFKENIAGVGLICSLIVLTQWRFLS 180

Query 181 FRKLMRWLLILVIALWFELDLYNISTCOPICAKWTRTEAGSKKKLSSGEGHMDLPDVV 240

Db 792 FRKLMRWLLILVIALWFELDLYNISTCOPICAKWTRTEAGSKKKLSSGEGHMDLPDVV 240

Query 241 ITSLPGSGAEILKQLFNNSDFLYIIRVPTAYIDIPETELEIDSFDVACEKVSIDRSCHF 300

Db 852 ITSLPGSGAEILKQLFNNSDFLYIIRVPTAYIDIPETELEIDSFDVACEKVSIDRSCHF 300

Query 241 ITSLPGSGAEILKQLFNNSDFLYIIRVPTAYIDIPETELEIDSFDVACEKVSIDRSCHF 300

Db 852 ITSLPGSGAEILKQLFNNSDFLYIIRVPTAYIDIPETELEIDSFDVACEKVSIDRSCHF 300

Query 241 ITSLPGSGAEILKQLFNNSDFLYIIRVPTAYIDIPETELEIDSFDVACEKVSIDRSCHF 300

Db 852 ITSLPGSGAEILKQLFNNSDFLYIIRVPTAYIDIPETELEIDSFDVACEKVSIDRSCHF 300

Query 301 RLJRGWLSQVDTKLHLQNHONTHLHEPNRSGKLAQYFAMNKDKRKFKFKEKRESLPEQRSONMK 360

Db 912 HLJRGWLSQVDTKLHLQNHONTHLHEPNRSGKLAQYFAMNKDKRKFKFKEKRESLPEQRSONMK 360

Query 361 AFDRDAEYTRALRRHLVYYPSSARVPLSLSGSNTKLHFFQEVLGASMRALYTVRDRAW 420

Db 972 AFDRDAEYTRALRRHLVYYPSSARVPLSLSGSNTKLHFFQEVLGASMRALYTVRDRAW 420

Query 421 IYSMLYNKSPLSYLSKVNYPHEALKLFKIEGGKGCNUNSGAYAEYEPRLKELSKSKSNAV 480

Db 1032 IYSMLYNKSPLSYLSKVNYPHEALKLFKIEGGKGCNUNSGAYAEYEPRLKELSKSKSNAV 480

Query 481 SLLSHLWLNATAALRINTDLLPSPYQVLFEDVHFQKTTERIIFAFLGIPSPASLNU 540

Db 1092 SLLSHLWLNATAALRINTDLLPSPYQVLFEDVHFQKTTERIIFAFLGIPSPASLNU 540

Query 541 ILFATSTNLFLYLPYEGETSPNTNVWONLPDEIKLLENCTWLMLRGYPKFMD 596

Db 1152 MLFATSTNLFLYLPYEGETSPNTNVWONLPDEIKLLENCTWLMLRGYPKFMD 596

RESULT 6

US-10-479-472A-4

; Sequence 4, Application US/10479472A

; GENERAL INFORMATION

; APPLICANT: DEL-FAVERO, JURGEN PETER LODE

; TITLE OF INVENTION: NOVEL BRAIN EXPRESSED GENE AND PROTEIN ASSOCIATED WITH

; TITLE OF INVENTION: BIPOLEAR DISORDER

; FILE REFERENCE: JAB-1711

; CURRENT APPLICATION NUMBER: US/10/479472A

; CURRENT FILING DATE: 2003-12-01

; PRIOR APPLICATION NUMBER: PCT/EP02/06316

; PRIOR FILING DATE: 2002-06-06

; PRIOR APPLICATION NUMBER: EP 01202214.1

; PRIOR FILING DATE: 2001-06-11

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn Ver. 3.2

; LENGTH: 1207

; SEQ ID NO 4

; TYPE: PRT

RESULT 7

US-09-384-429-71

; Sequence 71, Application US/09984429

; GENERAL INFORMATION

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 53 Human Secreted Proteins

; CURRENT APPLICATION NUMBER: US/09/984429

; FILE REFERENCE: P2018P2

; CURRENT FILING DATE: 2001-10-30

; PRIOR APPLICATION NUMBER: 60/244,591

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/288,143

; PRIOR FILING DATE: 1999-04-08

; PRIOR APPLICATION NUMBER: PCT/US98/21142

; PRIOR FILING DATE: 1998-10-08

; PRIOR APPLICATION NUMBER: 60/061,463

; PRIOR FILING DATE: 1997-10-09

; PRIOR APPLICATION NUMBER: 60/061,529

; PRIOR FILING DATE: 1997-10-09

; PRIOR APPLICATION NUMBER: 60/071,498

; PRIOR FILING DATE: 1997-10-09

; PRIOR APPLICATION NUMBER: 60/061,527

; PRIOR FILING DATE: 1997-10-09

; PRIOR APPLICATION NUMBER: 60/061,536

RESULT 9
US-10-021-660-128 Sequence 128, Application US/10021660
US-10-021-660-128 Publication No. US2003012926A1

GENERAL INFORMATION:
APPLICANT: BOS Biotechnology, Inc.
APPLICANT: Watson, Susan R.
APPLICANT: Murray, Richard
APPLICANT: Glynn, Richard
APPLICANT: Watson, Susan R.

TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis

FILE REFERENCE: 018501-000710B
CURRENT APPLICATION NUMBER: US/10/021,660
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: US 09/784,356
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: US 09/637,977
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 135
SOFTWARE: FastSEQ For Windows Version 3.0
SEQ ID NO 128
LENGTH: 411

TYPE: PRT
ORGANISM: Homo sapiens
US-10-021-660-128

Query Match 18.0% ; Score 568 ; DB 11; Length 125;
Best Local Similarity 100.0% ; Pred. No. 1.6e-46;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 MRWLLILVIALWFELDWDSTCSQPICAKTRTEAEGSKKSISSECHMOLPDVVITSL 244
Db 1 MRWLLILVIALWFELDWDSTCSQPICAKTRTEAEGSKKSISSECHMOLPDVVITSL 60

Query Match 18.0% ; Score 568 ; DB 11; Length 125;
Best Local Similarity 100.0% ; Pred. No. 1.6e-46;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 245 PGSGABILKQLFNSSDFLYRPTAYIDIPETELEIDSFVDAECK 291
Db 61 PGSGABILKQLFNSSDFLYRPTAYIDIPETELEIDSFVDAECK 107

RESULT 8
US-10-150-111-71 Sequence 71, Application US/10150111
; Publication No. US20030078366A1

GENERAL INFORMATION:
APPLICANT: Rubin et al.
TITLE OF INVENTION: Secreted Protein HPEAD48
FILE REFERENCE: P2018P1D1
CURRENT APPLICATION NUMBER: US/10/150,111
CURRENT FILING DATE: 2002-05-20
PRIOR APPLICATION NUMBER: 09/288,143
PRIOR FILING DATE: 1999-04-08
PRIOR APPLICATION NUMBER: PCT/US98/21142
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/061,463
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,529
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/071,498
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,527
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,536
PRIOR FILING DATE: 1997-10-09
PRIOR FILING DATE: 1997-10-09
PRIOR FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 219
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 71
LENGTH: 126
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: SITE
NAME/KEY: SITE
LOCATION: (126)
OTHER INFORMATION: Xaa equals stop translation
US-10-150-111-71

Query Match 18.0% ; Score 568 ; DB 14; Length 126;
Best Local Similarity 100.0% ; Pred. No. 1.6e-16;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 MRWLLILVIALWFELDWDSTCSQPICAKTRTEAEGSKKSISSECHMOLPDVVITSL 244
Db 1 MRWLLILVIALWFELDWDSTCSQPICAKTRTEAEGSKKSISSECHMOLPDVVITSL 60

RESULT 10
US-10-211-462-97 Sequence 97, Application US/10211462
; Sequence 97, Application US/10211462

; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: Methods of Screening for Angiogenesis, Compositions and Targeted Therapeutic Intervention for Mitochondrial Proteome
; FILE REFERENCE: 018501-0062001US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-211-462-97

Query Match Score 147; DB 15; Length 411;
Best Local Similarity 18.2%; Pred. No. 6.e-05;
Matches 89; Conservative 58; Mismatches 194; Gaps 16;

Qy 187 WILILVIALWIEL-----LDWSTC-----SQPICAWKTRTEAGSKKSLS 229
Db 5 WKAVLLALASIAQYTAIRTPAKSHTCPGLAEAGLAERL-----EESPPTFAYNLSR 59
Qy 187 WILILVIALWIEL-----LDWSTC-----LDWSTC-----LWVNLKLNQHDFPL 94
Db 5 WKVDIHSFHFLRGMIQSLVQDVTKHQLQNTLHHEPNRKGLAQYFAMNKDKCRKFRE 349
Qy 230 EGHHMDLPDVVITSLPGSAAELKQLEFFNSDFLYIRVPTAYIDIPETELEIDSFVDAE 289
Db 94 -----YHVQNTLIPRFQG-----YHVQNTLIPRFQG-----KSPADRR 115
Qy 290 WKVDIHSFHFLRGMIQSLVQDVTKHQLQNTLHHEPNRKGLAQYFAMNKDKCRKFRE 349
Db 94 -----YHVQNTLIPRFQG-----YHVQNTLIPRFQG-----KSPADRR 115
Qy 350 SLPEQRQSMKGAFDRDABYI-----RAILRHLYVYPSARPV-----LS 387
Db 60 KTH -----ILILATRSSSFVQLNHOHDVYFLPL-----93
Qy 290 WKVDIHSFHFLRGMIQSLVQDVTKHQLQNTLHHEPNRKGLAQYFAMNKDKCRKFRE 349
Db 94 -----YHVQNTLIPRFQG-----YHVQNTLIPRFQG-----KSPADRR 115
Qy 350 SLPEQRQSMKGAFDRDABYI-----RAILRHLYVYPSARPV-----LS 387
Db 116 MGASRDILRSLYDCDYFLENYIKPPVNHTRDRFRGASRVLCSPVDPGPADLV 175
Qy 388 LSGGSWTLKLHFFQEVIGA-----SMRALYTVRD 416
Db 176 LEEGDCVRKGCLNLNTVAEAACRSVHAIKVTVRPVENDRLALVEDPRLNKVIQLVRD 235
Qy 417 PRAWIYSMLYNEKSKPSLVSLSKVNPEHLAKLFKIEGGKGKGNLNSGYAFYEPERKELSK 476
Db 236 PRGLASRSETFRDY-----RLWRWYGTGR----KPYNLDVTQL-TTVCEDF 279
Qy 477 SNAVS--LISHMLANTAAALRINTDLIPTSYOLVKFTRIFAPLGPL- 533
Db 176 LEEGDCVRKGCLNLNTVAEAACRSVHAIKVTVRPVENDRLALVEDPRLNKVIQLVRD 235
Qy 417 PRAMYSMLNSKPSLVSLSKVNPEHLAKLFKIEGGKGKGNLNSGYAFYEPERKELSK 476
Db 236 PRGLASRSETFRDY-----RLWRWYGTGR----KPYNLDVTQL-TTVCEDF 279
Qy 534 -----SPASLNQILPATSTNLFLYFPEGEISOPTNTVWKQNLRDEIKLJENI 581
Db 280 SNSVSTGLMRPPWLKG-----KYMIVRYEDLARNMFKCTBEIYFLGFLGPLD 325
Qy 534 -----SPASLNQILPATSTNLFLYFPEGEISOPTNTVWKQNLRDEIKLJENI 581
Db 326 SHYARWIONNTRGDPFLGKHYGTVRN-----SAATAEKWRFLSYDIVAFAONA 375
Qy 582 CWTLMDRGY 591
Db 376 COQVLAQGY 385
RESULT 12
US-10-723-860-1544
; Sequence 1544, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions & Methods for Screening for Soft Tissue Sarcoma Modulators
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882-0193.NPUS01
; Sequence 395, Application US/10408765A
; Publication No. US20040101874A1

CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIORITY NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Mus musculus
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-723-860-1544

Query Match 4.7%; Score 147; DB 16; Length 411;
Best Local Similarity 18.2%; Pred. No. 6.4e-05; Gaps 16;
Matches 89; Conservative 58; Mismatches 149; Indels 194; Gaps 16;

Qy 187 WILLIVILALMFIEL-----LDVWSTC-----SOPICAKWTRTEAESGKSLSS 229
Db 5 WKAVILLALASIAQYTAIRTFATSKSFHICPGLAEGHLRLC---EESPTFAYNLSR 59

Qy 230 EGHMMDLPDVITSLPGSGAEILKOLFFNSDFLYTIVPFTAYDIPETELEIDSFDVACE 289
Db 60 RTH-----ILILATTRSGGSFVGOLFNOHDVYLFPFL-----93

Qy 290 WKVSDIRSGHFRLLRGWLQLSQVDTKLHQINLHEPNRQKLAQYFAMNKKRKFKERRE 349
Db 94 -----YHNTQNLTPRFQ-----KSPADRV 115

Qy 350 SLPBEQRSMQMGAFDRAEVT-----RALRHVVYYPSPRV-----LS 387
Db 116 MLGASRDLLSLYCDLYPFLENYIKPPPVNHTTDIPIRGASRVLCSRPVCDPPGPADIV 175

Qy 388 LSGGSWTLKHFQEVGLGA-----SMRALTYRD 416
Db 176 LEEGCVRKCGGLNLTVAEACRERSHVAIKTVRPEVNIDRLAVALPRLNKYTQYRD 235

Qy 417 PRANITYSMLYNSKPSLYSLKVNPEHLAKLFKIEGGKGKCNLNSGYAFAYEPLRKELSKSK 476
Db 236 PRGTIASRSRFDRY-----RURLWYGTGR---KEYNLNDYQL-TIVVCBEDF 279

Qy 477 SNAVS--LLEHLWINTAAALRINTDLLPTSYQLVKFEDIVHFPKQTTERIFALGIPL 533
Db 280 SNSVSTGLMPMPWLG-----KYMVLVRYEDLARNPMKKTBEIYGEJGIPID 325

Qy 534 -----SPASLNQILFATSTNLFLPYLYGEIEISPTNTNTWAKONLPRDEIKLIENTI 581
Db 326 SHVARWQNNTRGDPTLGKHYGTVRN-----SATAEKHRFLRSYDIVAFQAQNA 375

Qy 582 CWTLMDRIGY 591
Db 376 CQVLAQOLGY 385

RESULT 13
US-10-212-933-2
; Sequence 2, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadonatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takeshi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; TITLE OF INVENTION: DNA ENCODING THE SAME
; CURRENT APPLICATION NUMBER: US/10/212,933
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/09/263,023
; PRIOR FILING DATE: 1998-03-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-212-933-2
; Sequence 2, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadonatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takeshi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; TITLE OF INVENTION: DNA ENCODING THE SAME
; CURRENT APPLICATION NUMBER: TOYAM41.001AU5
; CURRENT FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US/10/212,933
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match 4.2%; Score 133; DB 14; Length 484;
Best Local Similarity 26.2%; Pred. No. 0.0019; Gaps 9;
Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 408 MRALYIVDRPRAIYMSLYNSKPSLY--SLKNVPE----HLAKLFKIBEGK-GKCNLNS 459
Db 279 LKVTHLVRDERA-VASSRISRSHGLIRESELQVRSRSDPRAHMRMPFLEAAGHKLGAKKEGV 337

Qy 460 GYAFFYEPURKELSKSKSNAVSLLSHWLANTAAALRINTDLLPTSYQLVKFEDIVHFPQ 519
Db 337 GGPDYHAL-----GAMEVICNSNAKTLQTLAQ-PDWLQGHYLVRYEDLVGDPV 386

Qy 520 KTTERRIFAFPLGIPLSPA---SUNQILPATSTNFYLPYEGEISPTN---TNWQKQNLP 571
Db 387 KTLLRVDVFGVLSPMEMEQFALMTSGSGSSSSKPKPV---VSARNTAQANAWRALT 441

RESULT 14
US-10-212-933-4
; Sequence 4, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takeshi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; TITLE OF INVENTION: DNA ENCODING THE SAME
; CURRENT APPLICATION NUMBER: TOYAM41.001AU5
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/10/212,933
; PRIOR FILING DATE: 1998-03-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match 4.2%; Score 133; DB 14; Length 484;
Best Local Similarity 26.2%; Pred. No. 0.0019; Gaps 9;
Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 408 MRALYIVDRPRAIYMSLYNSKPSLY--SLKNVPE----HLAKLFKIBEGK-GKCNLNS 459
Db 279 LKVTHLVRDERA-VASSRISRSHGLIRESELQVRSRSDPRAHMRMPFLEAAGHKLGAKKEGV 337

Qy 460 GYAFFYEPURKELSKSKSNAVSLLSHWLANTAAALRINTDLLPTSYQLVKFEDIVHFPQ 519

RESULT 15
 US-10-723-860-1409
 ; Sequence 1409, Application US/10723860
 ; Publication No. US2004023606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsburg, Wendy M.
 ; APPLICANT: Zlotnik, Albert
 ; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions & Methods for Screening for Soft Tissue Sarcoma Modulators
 ; FILE REFERENCE: 05882-0193.NPUS01
 ; CURRENT APPLICATION NUMBER: US/10/723, 860
 ; CURRENT FILING DATE: 2003-11-26
 ; PRIOR APPLICATION NUMBER: 60/429,739
 ; PRIOR FILING DATE: 2002-11-26
 ; NUMBER OF SEQ ID NOS: 8393
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1409
 ; LENGTH: 530
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-723-860-1409

Query Match Score 133; DB 16; Length 530;
 Best Local Similarity 4.2%; Pred. No. 0.0021;
 Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 408 MRAIYIVRDPRTYSMLYNKPSLY--SLKVNPE---HLAKLPXIEGGK-GKCNLNS 459
 Db 325 LKVHLVRDRA-VASSIRISRLGIRESLQVRSRDPRAHMPFLPAAGHLGAKKEGV 383
 Qy 460 GYAFFEYEPKRELKSKSNAVSLSLWLNATAAALRINTDLPTSOLVKFEDIVHFHQ 519
 Db 384 GGPADYHAL-----GAMETICNSMAKLTOTALQ-PPDWLQGHYLVRRYEDLVGDPV 433
 Qy 520 KTERIFAPLGIPLSPA---SUNQILPATSTNLFLPYEGEISPTN---TNWVKONLP 571
 Db 434 KTLLRVYDFVGLLVSPEMEQFALNMTSGSSSSKPFV----VSARNATQAAANAWRTALT 488
 Qy 572 RDEIKLIENTCWTLMDRUGY PK 593
 Db 489 FQQIKQVEEFYCOPMAVGYER 510

Search completed: June 23, 2005, 09:27:55
 Job time : 77.3449 secs

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 08:33:38 ; Search time 13.5183 Seconds
(without alignments)
4242.027 Million cell updates/sec

Title: US-10-697-828-9
Perfect score: 3147
Sequence: 1 MNRYNGAMMDVDAHYRMFW.....LIENICWTLMDRGYPKFMD 596

Scoring table: BLOSUM62 Gapopen 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : PIR_79:
1: pir1:
2: pir2:
3: pir3:
4: pir4:
5: pir5:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---|
| 1 | 133 | 4.2 | 484 | 2 JE0261 | N-acetylglucosamin N-acetylglucosamin |
| 2 | 113.5 | 3.6 | 484 | 2 JC7350 | hypothetical prote hypothetical prote |
| 3 | 109.5 | 3.5 | 1441 | 2 T00335 | rhotropic protein - hypothetical prote |
| 4 | 107 | 3.4 | 2401 | 2 T286756 | N-acetylglucosamin N-acetylglucosamin |
| 5 | 106.5 | 3.4 | 532 | 2 H96840 | hypothetical prote hypothetical prote |
| 6 | 106 | 3.4 | 486 | 2 JC7351 | hypothetical prote phytoene desaturase |
| 7 | 105.5 | 3.4 | 518 | 2 B83472 | probable membrane RNA-directed DNA P |
| 8 | 103 | 3.3 | 582 | 2 T46822 | DNA primase (EC 2.2.1.1) alpha-1-antitrypsin exopolyphosphatase |
| 9 | 102.5 | 3.3 | 689 | 2 S65901 | hypothetical prote hypothetical prote |
| 10 | 102 | 3.2 | 1333 | 2 S65812 | probable membrane RNA primase (EC 2.2.1.1) |
| 11 | 101.5 | 3.2 | 591 | 2 B97994 | chondroitin 6-sulf |
| 12 | 101 | 3.2 | 486 | 2 T28923 | probable mucin G2 |
| 13 | 100 | 3.2 | 1663 | 2 T28923 | probable membrane |
| 14 | 98 | 3.1 | 480 | 2 T20168 | hypothetical prote |
| 15 | 98 | 3.1 | 586 | 2 H95123 | ORF MSV089 probable |
| 16 | 97.5 | 3.1 | 834 | 2 T28250 | alpha-1-antitrypsin |
| 17 | 97 | 3.1 | 420 | 2 A28892 | conserved hypothetical |
| 18 | 97 | 3.1 | 561 | 1 G64246 | chondroitin 6-sulf |
| 19 | 96 | 3.1 | 458 | 2 A57397 | probable membrane |
| 20 | 96 | 3.1 | 1692 | 2 G01449 | hypothetical prote |
| 21 | 96 | 3.1 | 4910 | 2 B86294 | conserved hypothetical |
| 22 | 95 | 3.0 | 522 | 2 C90535 | hypothetical prote segment S5 protein |
| 23 | 95 | 3.0 | 963 | 2 1405 | hypothetical prote |
| 24 | 94 | 3.0 | 804 | 2 A39972 | hypothetical prote |
| 25 | 93.5 | 3.0 | 825 | 2 T32130 | hypothetical prote |
| 26 | 93.5 | 3.0 | 648 | 2 A81349 | probable ATP-depen |
| 27 | 93 | 3.0 | 676 | 2 A81349 | probable ATP-depen |
| 28 | 93 | 3.0 | 868 | 2 S65186 | NIP80 protein - ye |

ALIGNMENTS

RESULT 1
JE0261
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
C.Species: Homo sapiens (man)
C.Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 01-Mar-2002
C.Accession: JE0261
R.Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, Q.W.; Mitsuoka, T.
J. Biochem. 124, 670-678, 1998
A.Title: Human N-Acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of chondroitin 6-sulfotransferase
A.Reference number: JB0261; PMID:98291845; PMID:9722682
A.Molecule type: mRNA
A.Residues: 1-484 <UCH1>
A.Cross-references: DDBJ:AB0144679
C.Comment: This protein catalyzes the transfer of sulfur of sulfate from 3'-phosphoadenosine 5'-phosphate to the 6-O-sulfate group of chondroitin 6-sulfotransferase
C.Superfamily: chondroitin 6-sulfotransferase
C.Keywords: sulfotransferase

Query Match Score 133; DB 2; Length 484;
Best Local Similarity 26.2%; Pred. No. 0.023; Mismatches 36; Indels 32; Gaps 9;
Matches 53; Conservative 36; MisMatches 81;

Qy 408 MRLAYIVDPRAVTYSLYNSKPSLY--SLKLVPE----HLAKLFLIEGGK-GKCNLNS 459
Db 279 LKVTHLVDPRA-VASSRIRSRLGLIRESLQVRSRSDPRAHMPFLAEGHKLGAKKEGV 337
Qy 460 GYAFEEYEPKRLKLSKSNSNAVSSLHMLANTAAALRINTDLUPTSYQLVKFEDIVHFFQ 519
Db 338 GGPDADYHAL-----GAMETVTCNSNPKTLQTALQ-PPDWLQGHYLVRYEPLVGEPV 387
Qy 520 KITPERIAFLGLIPUSA---SUNQILPATSNLFLYVEGETSPNTN---TNTWKQNLP 571
Db 388 KTLLRVYDFGVLLSPMEQFALNMTSGSSSSKPFV----VSARNATQAANAWRTALT 442
Qy 572 RDEIKLTIENICWTLMDRGYPK 593
Db 443 FQQIQKQVEEFCYQPMAVGYER 464

RESULT 2
JC7350
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse
C.Species: Mus musculus (house mouse)
C.Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
C.Accession: JC7350
R.Uchimura, K.; Fasakhany, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurokawa, N.; Biotech. Biophys. Res. Commun. 274, 291-296, 2000
A.Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
A.Reference number: JC7350; PMID:20374462; PMID:1091333
A.Accession: JC7350

A; Molecule type: mRNA
 A; Residues: 1-484 <UCH>
 A; Cross-references: UNIPROT:Q99N80; UNIPROT:Q9EP78; DDBJ:AB040710
 C; Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in
 C; Keywords: glycolysis; sulfotransferase; transmembrane protein

Query Match Score 113.5; DB 2; Length 484;
 Best Local Similarity 18.8%; Pred. No. 0.74%; Indels 107; Gaps 18;
 Matches 81; Conservative 72; Mismatches 171;

| | | | |
|----|-----|--|-----|
| Qy | 222 | GSKSLSSSEGHHM-DLDPVITSLPGSGAEILKLQFLPNSDFLYIRVPTAYIDIPETEL | 279 |
| Db | 83 | GSPNLNSAVGEATQEKQHIVYHATWRTGSFGLGELPNQHPDVFLYEPWMHL | 135 |
| Qy | 280 | EIDSFVDAECKWYKSDIRGCHFLLRGWQLSLOVD--TKLHQINIHHEPPNGKLQYFA | 336 |
| Db | 136 | WQA-LYFGDAESQIQLRDLRSFLRCDESVLRLYAQPGDGERAPDSA | 183 |
| Qy | 337 | MNKDKRKRFKRR-----ESLPEQRQSQMGAFDRDAE-----YIRALRHLVY | 379 |
| Db | 184 | -NLTTAMEFRWRNTKVICSPPLCPAAPPARADIGLVRAKACSTCPVSLRAEAERKY | 242 |
| Qy | 380 | PSA-RPVPLSLSGGSSWTLKLHFFQEVL-----GASMRLAYIVRDPRAWIYSMLYNNSKP | 430 |
| Db | 243 | PVVVKDVRLDLG-----VLVPLLRDGPGLNLKVQVLERDRA-VHNSRLKSQR | 290 |
| Qy | 431 | SL-----YSLKNVPEHLAKLKLIEGSKGKCNLNQGYAFAYTEPLRKELSGKSNSAVSL | 483 |
| Db | 291 | GILLRESIQVLRTQRGDIFHRVLLAHG---WDARPGQARALPSAPRADPFLTSALEVI | 346 |
| Qy | 484 | SHWLANTAAALRINTDLPT-----SYOLVKEDIVHFPQKTERIAFAFLGILPLS | 534 |
| Db | 347 | CEANL-----RDLLTRGAPANLRLRRLRRTEDVWQPQAQLRLLRFSGL-RT | 395 |
| Qy | 535 | PASLNQILPATSTNLFLPYEGEGISPTN-----TNWKQNLPRDEIKLLENICWTL | 585 |
| Db | 396 | LAALDAFAFMTRGSAY---GADRPFHLSARDAREAVHMRERLSEOQVROVETACAPA | 451 |
| Qy | 586 | MDBLGYPKFMD 596 | |
| Db | 452 | MRLLAYPRSGD 462 | |

RESULT 4
 T28676
 Rhoptry protein - Plasmodium yoelii (fragment)

C; Species: Plasmodium yoelii
 C; Accession: T28676; A45521
 C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 R; Sinha, K.A.; Keen, J.K.; Ogn, S.A.; Holder, A.A.
 R; Sinha, K.A.; Keen, J.K.; Ogn, S.A.; Holder, A.A.
 R; Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
 R; Biochem. Parasitol. 76, 329-332, 1996
 R; Biochem. Parasitol. 42, 241-246, 1990
 R; Reference number: 205057; MUID:97077455; PMID:8920022
 R; Accession: T28676
 R; Status: Preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: DNA
 A; Cross-references: 1-2401 <KEB>
 A; Residues: 2260-2401 <KEB>
 A; Status: preliminary
 A; Molecule type: DNA
 A; Cross-references: GB:M34281

Query Match Score 107; DB 2; Length 2401;
 Best Local Similarity 20.8%; Pred. No. 21;
 Matches 62; Mismatches 129; Indels 106; Gaps 17;

| | | | |
|----|-----|---|-----|
| Qy | 66 | FYGPYINVSSCRFDSSNPGLQISL----NNNTTENVIVSIYDHNLRTRNYLGGFA | 121 |
| Db | 451 | FFGTGLDIFSCHAVDNPNPKPHSELEMLROYGPVNPEPLQKV-----AAGELR | 499 |
| Qy | 122 | SVADQGQTTRFGIYQTOAIVKPVYRDRITFPGFKENIAVGJLICISLIVLTFQWRFLSF | 181 |

RESULT 3
 T00335
 Hypothetical protein KIAA0564 - human (fragment)
 C; Species: Homo sapiens (man)
 C; Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
 A; Accession: T00335
 R; Nagase, T.; Isekiwa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
 DNA Res. 5, 31-39, 1998
 A; Title: Prediction of the coding sequences of unidentified human genes. IX. The complete
 A; Reference number: 9628581
 A; Accession: T214086; MUID:9820545; PMID:9820545
 A; Status: Preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: mRNA
 A; Cross-references: UNIPROT:O60310; EMBL:AB011136; NID:93043651; PIDN:BAA22490-1; PIDN:93
 A; Experimental source: brain; clone HH1811
 C; Genetics:
 A; Note: KIAA0564

Query Match Score 109.5; DB 2; Length 1441;
 Best Local Similarity 19.1%; Pred. No. 6.8;
 Matches 115; Conservative 93; Mismatches 219; Indels 175; Gaps 32;

| | | | |
|----|-----|---|-----|
| Qy | 66 | FYGPYINVSSCRFDSSNPGLQISL----NNNTTENVIVSIYDHNLRTRNYLGGFA | 121 |
| Db | 451 | FFGTGLDIFSCHAVDNPNPKPHSELEMLROYGPVNPEPLQKV-----AAGELR | 499 |
| Qy | 122 | SVADQGQTTRFGIYQTOAIVKPVYRDRITFPGFKENIAVGJLICISLIVLTFQWRFLSF | 181 |

Query Match Score 107; DB 2; Length 2401;
 Best Local Similarity 20.8%; Pred. No. 21;
 Matches 62; Mismatches 129; Indels 106; Gaps 17;

| | | | |
|----|-----|---|-----|
| Qy | 289 | EWKVSDIISGHFRLRGWL-OQLVDTKLHQ-NIHLHEPNRGLQLAQYFA----- | 336 |
| Db | 929 | EBKISDIRKNSLKIQDFSESSYINIDKKEBLEKVKLESQNNNTDINQYUSKIENTINYILK | 988 |
| Qy | 337 | MNKDKRKRFKRR-----ESLPEQRQSQMGAFDRDAE-----YIRALRHLVY | 379 |

RESULT 5
hypothetical protein F23A5.18 [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)
C;Accession: F96840
R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federepiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chinn, C.W.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Conway, A.B.; Conway, T.H.; Dewar, K.; Rizzou, M.; Rooney, T.; Rowley, D.; Sakano, H.; Maiti, R.; Marzaioli, R.; Authors: Saalberg, C.M.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A;Reference number: A86441; PMID:21016719; PMID:1113012
A;Accession: F96840
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-532 <STO>
A;Cross-references: UNIPROT:Q9SA4H; GB:AE005173; NID:96503294; PIDN:AAF14670-1; GSDB:GN
C;Genetics:
A;Gene: F23A5.18
C;Map position: 1
C;Superfamily: natural resistance-associated macrophage protein 1

Query Match 3.4%; Score 106.5; DB 2; Length 532;
Best Local Similarity 19.7%; Pred. No. 2.9; -Mismatches 90; -Indels 210; -Gaps 33;

Db 23 NSPLIENSDNQIIYSEKKSMKNEF -FAYLGPGFLVSIATIDPGNFETDIQAGAHKYKL 80
Qy 71 ---INVSSCRFDSSNPGLQISLNNTTEHVSVITVDYHNLKTRFNYLGGFFASVADQ 127
Db 81 LWILIVASC-----AALVIQSLAANLGVVTGHIAE-----OCRAEYS 118
Db 238 QITRFELGTQDIAVKPVYRHD--RII---FPFGFKENIAV--GLIL--CISVILITFQWRY 178
Qy 119 KVPNPMLWVVAEIAVACDPEVIGTAFLANMLFSIIPVNWCVLTLGSLTLLAQKYGWV 178
Db 179 LSFRKLMRWWLILVIALWFELLDWTSCTCOPICAKWTRAEAGSKSISSEGHHMDLP 238
Qy 239 -----VVATSLSPSGCAEYL-----KQLFNSSDFLYIRVPTAYIDIPETELE 280

Db 989 LNKKIKIIDKVKEYTDEIEKNNKKINAELUSNEKEITQLKENSLSKECOSKKIKSTIDDNY 1048
Qy 366 ---AEYTRALRRHLV-----YPPSAPRV--LSSGSWSLTKLUFFQEVLGASHRA 410
Db 1049 VSEC1KNTNLTQVTKVNEKNINTYFKAENEYONVSLN-----FNNIEADFTKS 1098
Qy 411 LYIRDPRAMWYMSLYNSKPSLISLNKVPDEPHLAK-LFKEGGKG---KCN----456
Db 1099 QYLNNIKON-----NGTNNTDYNKIKELKEHCKKSNSNVYKDEAGKNTQESTKCNKELFEKYE 1152
Qy 457 -----LNSCYAFYEPLRKELSKSKSNASVLLSHWLWN---TAAARLINTDLPLTYSQ 507
Db 1153 QEVTLNKKYAVE--LNRKFDKTKNYSQQIKEIKDAHNTFTSQADKSEKMMNEIKNE 1209
Qy 508 LVKEDIVHPPQKTTTERIFPLFLGIPPLSPASLNQLFATSTNLFLYLPYE--GEISPTNTNV 565
Db 1210 QIRTEDEVAKNNNSNKN---AIDLQISVEBF-KIKFLKIKDLRTKSDDOLEKTDIETKI 1265
Qy 566 WKQNLRPDETKLLEN 580
Db 1266 SNLSIDTQETKLIEN 1280

RESULT 6
hypothetical protein F23A5.18 [imported] - *Arabidopsis thaliana*

JC7351
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - human
C;Species: Homo sapiens (man)
C;Accession: JC7351
C;Sequence revision 08-Sep-2000 #text_change 09-Jul-2004
R;Uchimura, K.; Pasakiany, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurokawa, N.; Biochem. Biophys. Res. Commun. 274, 291-296, 2000
A;Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a r
A;Reference number: JC7350; PMID:20374462; PMID:10913333
A;Accession: JC7351
A;Residues: 1-486 <UCH>
A;Cross-references: UNIPROT:0755667; DDBJ:AB040711
C;Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in C;Keywords: glycosylation; sulfate transport; sulfotransferase; transmembrane protein

Query Match 3.4%; Score 106; DB 2; Length 486;
Best Local Similarity 20.6%; Pred. No. 2.8; -Mismatches 136; -Indels 154; -Gaps 18;
Matches 87; Conservative 45; -

Db 179 LSFRKLMRWWLILVIALWFELLDWTSCTCOPICAKWTRAEAGSKSISSEGHHMDLP 237
Qy 187 LTAAALFRRTKNTVLT-----CSPPPLCPGAPPARAEGVGLVEDTACE---RSSCP 230
Db 238 DVVFTSLPSSGAETLKQPFNNSSDFLYIRVPTAVIDIPETELEDSVIDACEWKVSDIRS 297
Qy 231 PVAIRALE--AECRK-----YPVVI-----KDVYR- 253
Db 298 GHFRFLRGFLQSLYQDTKLHQLNTHLHEPNRKLAQYFAMNKDKKKRKFGRRESLPEQRSO 357
Db 254 ---LDLGVLPLRDGPGLNKLKVQLFRDPA---VENSRLKSGRQLLRESQVLRTR 305
Qy 358 MKGAFDRAFTAEYTRALRHLYYUYPSSARPVLSGSSWSLTKLHFFQEVLGASMRALYIVRDP 417
Db 306 QRG--DR---FHRVLLAHGV---GARP-----GGOSRALDPA--P 335
Qy 418 RAWYSMLNSKPSLISLNKVPDEPHLAKUFKIEGGKGKCNLNNSGYAFEYBLRKELKSKS 477
Db 336 RADDF--LT 342

| | | | |
|----|---|-----|---|
| Qy | 534 SPASLNQIIFAT--STNLFLYLPYEGEISPTNTNWKQNLPDEIKLLENICWTLMRDGY | 591 | |
| Db | 402 DAFLLANMTGGAYGADRPHL--SARDAREAVHAWRERLSREQRVQEACAPAMRLAY | 459 | |
| Qy | 592 PK 593 | | |
| Db | 460 PR 461 | | |
| | | | RESULT 8 |
| | | | T46822 |
| | | | phytoene desaturase [EC 1.3.-.-] [validated] - Xanthophyllomyces dendrorhous |
| | | | C.Species: Xanthophyllomyces dendrorhous |
| | | | C.Date: 17-Mar-2000 #sequence_change 17-Mar-2000 #text_change 09-Jul-2004 |
| | | | C.Accession: T46822 |
| | | | C.Species: Preliminary; translated from GB/EMBL/DDJB |
| | | | R.Verdoes, J.C.; Misawa, N.; van Ooyen, A.J.J. |
| | | | Biotechnol. Bioeng. 63, 750-755, 1999 |
| | | | A.Title: Cloning and characterization of the astaxanthin biosynthetic gene encoding phytc |
| | | | A.Reference number: 224099; MUID: 99326230; PMID: 10397832 |
| | | | A.Accession: T46822 |
| | | | A.Status: Preliminary; translated from GB/EMBL/DDJB |
| | | | A.Molecule type: DNA |
| | | | A.Residues: 1-582 <VER> |
| | | | A.Cross-references: UNIPROT:O13506; EMBL:Y15007; NID:g2505954; PIDN:g2505954; PID:9250: |
| | | | A.Experiment source: strain CBS 6938 |
| | | | A.Genotype: crtI |
| | | | A:Introns: 16/1; 40/3; 56/3; 60/3; 91/1; 101/3; 136/3; 175/3; 233/2; 351/1; 401/3 |
| | | | C.Function: catalyzes the conversion of prephytene pyrophosphate to phytoene [validate] |
| | | | C.Superfamily: phytoene dehydrogenase |
| | | | C.Keywords: Oxidoreductase |
| | | | Query Match 3..31; Score 103; DB 2; Length 582; |
| | | | Best Local Similarity 19.5%; Pred. No. 6.2; |
| | | | Matches 73; Conservative 43; Mismatches 123; Indels 136; Gaps 14; |
| Qy | 322 HLHEPNRGKLAQYFANKDKD--KRFKFRKRESLPEQRSQMKGAFDRDAEYIRALRRH--- | 375 | |
| Db | 98 HPHD-----ETTFETSTDMAILKREVERFEG-----KDGFEDRLTSFQEABHRYEL 143 | | |
| Qy | 376 LIV-----YPSARPVYSLSSGSWTLKHFEDVGLSMRALYIVDPRPNTYSMLY--- | 426 | |
| Db | 144 AVHVVLQKNPQPFPAAFRLQFQGQILAHLPFFESIWRVCYRPFKTDRLRRVFSFVMMGQ 203 | | |
| Qy | 427 --NSKPLSLSK-----NYPHEHLAQLFKLEGGSKCNLNNSGYAFFYEP 467 | | |
| Db | 204 SPYSPARTYSLQQYTELTETGIVWYPRGGFWQVENTLQGIVRNPNPSAKFNNA-----P 256 | | |
| Qy | 468 LRKE-LSKSKSNAVSLLSHMLANTAAALRINTDLLETSYQLVK----- | | |
| Db | 257 VSQVLLSPAKDADTGVLESGREENHADLYVASEHLPDDARNKIGOLGEYERS 316 | | |
| Qy | 511 QKTRTRIAFLGIPISPA 536 | | |
| Db | 317 WADLVGKKLKGSCSSLSFYMSMDRIVDGLGGHHNIFLAEDFKGSFTTIEBFLGPDP 376 | | |
| Qy | 537 SLNOILPATSTNLFLYLYEGEISPTNTNWKQNLP-----PRDEIKLHEN--- | 580 | |
| Db | 377 -----FVNVPSRIDPSAPEGKDAIVILVPCGHIDASNPDYNKLVARARK 423 | | |
| Qy | 581 -ICWTLMDRLGYPKF 594 | | |
| Db | 424 FVQTQSAKLGPDF 438 | | |
| | | | RESULT 9 |
| | | | S45901 |
| | | | probable membrane protein YBR043C - Yeast (Saccharomyces cerevisiae) |
| | | | N Alternate names: hypothetical protein YBR043C |
| | | | C.Species: Saccharomyces cerevisiae |
| | | | C.Date: 26-Aug-1994 #sequence_change 09-Sep-1994 #text_change 09-Jul-2004 |
| | | | C.Accession: S45901 |
| | | | R.Andre, B.; Cziepluch, C.; Hein, C.; Jauniaux, J.C.; Urrestaazu, A.; Vissers, S. |
| | | | submitted to the Protein Sequence Database, August 1994 |
| | | | A.Reference number: S45893 |
| | | | A.Accession: S45901 |
| | | | A.Molecule type: DNA |
| | | | A.Residues: 1-689 <AND> |
| | | | A.Cross-references: UNIPROT:P38227; EMBL:Z35912; NID:g536267; PIDN:CAA04985..1; PID:9516262 |
| Qy | 546 STNLFLYLPYEGEISPTNTNWKQNLPDEIKLLENICWTLMRDGY 591 | | |
| Db | 477 LRNNRPLPDF-----EPRAWSKGLAPDEQELVANIVCDVAERFH 516 | | |

A;Experimental source: strain S288C
 C;Genetics:
 A;Cross references: SGD:S0000247
 A;Gene: MIPS:YBR043C
 A;Map position: 2R
 C;Superfamily: Yeast probable membrane protein YBR043C
 C;Keywords: transmembrane protein
 F:140-161/Domain: transmembrane #status predicted <TM1>
 F:177-193/Domain: transmembrane #status predicted <TM3>
 F:23-256/Domain: transmembrane #status predicted <TM1>
 F:265-287/Domain: transmembrane #status predicted <TM5>
 F:416-493/Domain: transmembrane #status predicted <TM6>
 F:511-529/Domain: transmembrane #status predicted <TM7>
 F:555-579/Domain: transmembrane #status predicted <TM8>
 F:587-603/Domain: transmembrane #status predicted <TM9>
 F:619-645/Domain: transmembrane #status predicted <TM10>
 F:648-670/Domain: transmembrane #status predicted <TM11>

Query Match Score 102.5; DB 2; Length 689;
 Best Local Similarity 21.8%; Pred. No. 8.6;
 Matches 78; Conservative 56; Mismatches 139; Indels 85; Gaps 14;

Qy 167 SLVLTFOFRFLSFRKLMRWILVIALWFIELL---DWSTCSQPICAK----- 214
 Db 253 SLLVNRWPMRSTQEMVILGNCNVILLLTLPETRKQDSKGAAQIILERRIQVDNNER 312

Qy 215 ----WTRPAAEKSK----KSLSSEGH-----MQLPDVYTTSPLPG----SGAEI 251
 Db 313 GEIQEDYQRGEDDTDRIENQATLSTEKHINVGGEVRDQSDLDLESHSSNPNTYDGRAGETQ 372

Qy 252 LKQLPFNNSDFYLIRVPTAYIDIPPETELEIDSFDACEWKVSDRSRGPRLRWLQSILV 311
 Db 373 LQRIVTEASSLY-----BYQLD-DSGIDATAAQTRRSTDKLARSIRENSL 420

Qy 312 Q--DTPLKLHQNLTHHEPNRQKLAQYFANMKDKKKFKRRESLPEORSONQKGAFRDAEY- 368
 Db 421 RKLQTNLLEEYKKVLSNGGBIA-----PKQVAVRKWDTFVVF 461

Qy 369 IRALER-HLYTYPSPARVPLSSGSWTLKHFQEVLGASMRALXYTIVRPRAWIYSMLYN 427
 Db 462 IKPLKSLHFLJEPYPPVALATPSAIS-----FSTVYFTMVTVEKYKSSRP-----YN 507

Qy 428 SKPSLYSKVNPHEIYLAKLF-KLEGKGKCNLNSGYAFEYPLRKELSKSNSAVSLLS 484
 Db 508 FKPLYIGLMLIPNSUTYFFASIYGGRRWDMILKRYKERGILAPEARIWNNTVIS 565

RESULT 10
 S65812 Directed DNA polymerase (EC 2.7.7.49) (clone DW1) - Arabidopsis thaliana retrotranscriptase
 N;Alternate names: reverse transcriptase
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 20-Jul-1996 #sequence_revision 27-Feb-1997 #text_change 09-Sep-1997
 C;Accession: S65812; S65814
 R; Wright, D.A.; Ke, N.; Smalle, J.; Hauge, B.; Goodman, H.M.; Voytas, D.P.
 R;Submitted to the EMBL Data Library, October 1995
 A;Description: Multiple non-LTR retrotransposons in the genome of *Arabidopsis thaliana*.
 A;Reference number: S65811
 A;Accession: S65812
 A;Molecule type: DNA
 A;Residues: 1-1333 <WRI>
 A;Cross-references: EMBL:L47193; NID:976276; PID:9976278
 R; Wright, D.A.; Ke, N.; Smalle, J.; Hauge, B.M.; Goodman, H.M.; Voytas, D.P.
 Genes 142, 569-578, 1996
 A;Title: Multiple non-LTR retrotransposons in the genome of *Arabidopsis thaliana*.
 A;Reference number: S65813; MUID:97005555; PMID:8852834
 A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 448-648; 680-696; 714-723; 747-757; 1086-1103 <WRN>
 A;Cross-references: EMBL:L47193
 C;Genetics:

A;Mobile element: retrotransposon 'tail-1'
 C;Keywords: nucleotidyltransferase; reverse transcriptase
 Query Match Score 102; DB 2; Length 1333;
 Best Local Similarity 17.8%; Pred. No. 23;
 Matches 100; Conservative 67; Mismatches 156; Indels 238; Gaps 29;
 Qy 104 DY--HNLLKTRENYLGFFGASVADGGQTTRFGJGTQAIYKPVHRDRLIFPPGFKFNIAVG 161
 Db 768 DYIKDRDLTRLS---GWFARTLSMG-----GKETLTK----- 796
 Qy 162 LILCISLVLTQFWRPFY-LSPRKLMRWWLILVIA---LWFIELDVSTCSQPICAKT 216
 Db 797 ---AFALALL---FVAMSCFLKTTCVNNTSAMDFW-----WNALIEHKRKTWHV 841
 Qy 217 RTEAEGSKSLSSEGHINDLPDV-----ITSLPGS-GAEILKQLQELPNSDF 262
 Db 842 SCE---KNCLSKENGUGFRDIESPNQALLAQAWRLLQFPNSLARFPKSRYDEEDF 897
 Qy 263 LYIRVPTAYIDIPETELEIDSFDACEWKVSDIRSGHPLLRG-----WLQS 309
 Best Local Similarity 21.8%; Pred. No. 8.6;
 Matches 78; Conservative 56; Mismatches 139; Indels 85; Gaps 14;
 Db 898 L-----DAELKA---TPSYAVR--SILHGRDLIKGFRKIVNGSSTSVMMDP 940
 Qy 310 LVQDPTKLHLQNLHILHEPORGKLAQYFANKDKK-----RKPKRKRESLPEQSQMKGKA 361
 Db 941 WIID-----NDP-RLPLQKHSFSYNDLRVHDLLINVEDCRDRDLEE----- 981
 Qy 362 FDRDAEYTRALRHLRVYTP-----ARPVLSSGWSWTLKHFQEVLGASHRALYIV 414
 Db 982 -----LFPADIEITLVRNPVNSHDD----- 1002
 Qy 415 RDPRAWIYSMLYNKSPSLYSLKVNPEHLAKLFKLEGKGKCNLNSGYAFEYPLRKEL-- 472
 Db 1003 ----FWW---LHSKSGEYSVK-----SGWLAFTQTNKPELIR 1033
 Qy 473 -SKRSKNAVSLSLHLLWANTAAL---RINTDLPLPTSYQLVK-----F 511
 Db 1034 EARVQPTSTNLKEKIWSLUTSPKIKFLWRILSSALPVAYQIIRRGMPLDPRCCQVCGBEG 1093
 Qy 512 EDITH--FQPKTQTERIATLGLPISPASt--NQILFATSTNLFLYFPEGYGISSPNTNWNWQ 568
 Db 1094 ESINHVLFCSLARQWVALSGVPTSQFGQNNSIPIANIQYLLELKKGKLLPEQIKKSWPW 1153
 Qy 569 NLPRDEIKLIENICWTLMDRL 589
 Db 1154 VLWR-----LWKNRDQL 1165

RESULT 11
 B97994
 DNA primase (EC 2.7.7.-) [imported] - *Streptococcus pneumoniae* (strain R6)
 C;Species: *Streptococcus pneumoniae*
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C;Accession: B97994
 R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczaik, L.; Burgett, S.; Dehoff, B.S.; Et
 al., R.; LeBlanc, D.J.; Lee, J.; Lefkowitz, B.J.; Lu, J.; Matsushima, P.; McAhren, S.; Mc
 J.; Bacteriol. 183, 5709-5717, 2001
 A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaszkunas, S.R.;
 A;Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
 A;Reference number: A97872; PMID:21429245; PMID:11544234
 A;Accession: B97994
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Cross-references: 1-591 <KUR>
 A;Cross-references: UNIPROT:Q8DPW1; GB:AE007317; PID:AAK9782.1; PID:915458592; GSPDB:Gr
 C;Genetics:
 A;Gene: anaG
 C;Superfamily: DNA primase
 C;Keywords: nucleotidyltransferase

Query Match Score 101.5; DB 2; Length 591;
 Best Local Similarity 19.6%; Pred. No. 8.3;

| | | | | | | | | | |
|--|-------------|--|--------------------------|----------------------------------|------------------------------|------------------------|-----------------------|--------------------------|---|
| Matches | 110; | Conservative | 79; | Mismatches | 206; | Indels | 165; | Gaps | 24; |
| Qy | 87 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | | | | | | |
| Db | 10 | QVIEEIKONANIVEVIGDVTSQLRAGRNTYGLCPFHGXKTPSFSVVEDQFYHCFGCGRS | 69 | | | | | | |
| Qy | 138 | AIVPKVRHDRIIFPGFKPENIYAVGLICLSSLVLTQWRFLYSLRKLWILLVIAWF | 197 | | | | | | |
| Db | 70 | GDV-----FKE-----EEFYQ-----GTF | 84 | | | | | | |
| Qy | 198 | IELDVWSTCSQPICAKWTRTEAGSKSLSSE----GHMHDLPD-----VVIT | 242 | | | | | | |
| Db | 85 | MEAQV-----ILGQRVGEVE--KPLYSEOKPASHQALQYDMHDAKFYHAINT | 133 | | | | | | |
| Qy | 243 | SPLSGA-----BILKOLFFN----SSDFLYIRPTAY--IDIPETEL---EI | 281 | | | | | | |
| Db | 134 | TNGBEARNLYXORGTLTDPLKHFWIGLAPERNLYQLRSQYREEDLDSLGFYLSDA | 193 | | | | | | |
| Qy | 282 | DSEVFACE-----WKVSDIRSGHzFRLLRGWLQSLVQDTKLHLQN | 320 | | | | | | |
| Db | 194 | NQFVDTFHNRIMPFNLNDQGVIAFSGRLWQKUDSQTSKYNSR---STVIFNKSYELYH | 250 | | | | | | |
| Qy | 321 | IHPHEPNRGLQAQYFAMMKDKCKKFKRREPSLPEORSMQGAFRD-AEYTRALRRHLYYY | 379 | | | | | | |
| Db | 251 | MDRAKRSSGKASETYLMGFMVDITYAAYRAGIENAVASNGTALSREHVEHLKRTKLV- | 308 | | | | | | |
| Qy | 380 | PSARPVLSSGSWTLKLFHFQEYLG--ASMRALYIVRDPRANTYSMLYN SKPSL SLSKRN | 437 | | | | | | |
| Db | 309 | -----LIVYDGDKAGQATLKADEIGMPVQVSMPLNLDPDEYLQRN | 351 | | | | | | |
| Qy | 438 | VPEHLAKLFKIEGGKGNLNSGYAFYEYEP-----LRLKLSKSNSNAVSLSHIWMLANTA | 492 | | | | | | |
| Db | 352 | SPEDLAYLT----KTRISPFLPYTHQYKPGENGNLQAOIEFLKEIAPLIVQEKSIAQN | 407 | | | | | | |
| Qy | 493 | AAIRINTDLPTTYQLVKPEDIYHFP----QKTERIFAPLGFIPPLSPASLNQILFATST | 547 | | | | | | |
| Db | 408 | SYHILADSL-ASFDYTQIEQIYNEQRVQVRQNRMER1SRPTPTMPVTKQLSAMRAEA | 466 | | | | | | |
| Qy | 548 | NLFYLPYEGEISPTNTNYWK | 567 | | | | | | |
| Db | 467 | HLLYRMME--SPLVLDNYR | 483 | | | | | | |
| RESULT 12 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 13 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 14 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 15 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 16 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 17 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 18 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 19 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 20 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 21 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 22 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 23 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-668, | 2000 | A; GenBank accession: H81377 | A; Status: Preliminary | A; Molecule type: DNA | A; Residues: 1-486 <PAR> | Query Match Score 3.2%; Best Local Similarity 21.4%; Prod. No. 6.9; Mismatches 56; Indels 106; Gaps 18; |
| C; Cross-references: UNIPROT:Q9P1EB; GB: ALA39075; GB: AL111168; NID: 96967817; PIDN: CAB7419 | | | | | | | | | |
| A; Experimental source: serotype O2, strain NCTC 11168 | | | | | | | | | |
| C; Genetics: A; Gene: CJ053_c | | | | | | | | | |
| C; Keywords: hydrolase | | | | | | | | | |
| RESULT 24 | | | | | | | | | |
| H81377 | 80 | QISLNVNNTTEHVSIVTVDYHNL-KTRFNYLG----FGG-----FASVADQGQTTRFLGLSQ | 137 | C; Species: Campylobacter jejuni | Imported | - | Campylobacter jejuni | (strain NCTC | 268 |
| C; Accession: H81377 | 31-Mar-2000 | #sequence_revision 31-Mar-2000 | #text_change 09-Jul-2004 | | | | | | |
| R; Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell, B.; Reference number: AB1250; PMID: 10688204 | Nature | 403, | 665-6 | | | | | | |

A; Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A; Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
 A; Reference number: A55000 ; PMID:21357209 ; PMID:1163916

A; Accession: H95123

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-586 <KUR>

A; Cross-references: UNIPROT:Q970X2; GB:AB005672; PIDN:AAK75185.1; PID:914972547; GSPDB:GR

A; Experimental source: strain TIGR4

C; Generics:

A; Gene: SP1072

Query Match 3.1%; Score 98; DB 2; Length 586;

Best Local Similarity 18.8%; Prod. No. 15; Mismatches 83; Indels 210; Gaps 27; Matches 117; Conservative 83;

Qy 87 QISLUNVNNTTEHVVSIIVTDDYHNL-KTRFNYLG---FGG----FASVADQGQITRFGLGTO 137
Db 5 QVBEIKNNANIVEVIGIVVISLQAGRNYGLCPFHGXTPSPNVEDRQFYHFGCGRS 64Qy 138 AIYKPVHRDRIIFFFFKFNIAVGLLICSLVLTFRQWRFYLSPKLMRWWLILVIALMP 197
Db 65 GDV---FKEF-----IEYYQ-----GVPF 79Qy 198 IELLDWWSCTCSOPICAVTRTEAGSKSLS---SEGHHMDLPD---VVITS 243
Db 80 IEAVQILG-----QRVGIEVEKPLYSBQKSAPHQALIDMHEDAKFYHAIIMTT 129Qy 244 LPSSGA-----EYLKOLFFN---SSDFLYIRVPTAY--IDIPETEL---EID 282
Db 130 TMGBEARNYLYQRLGTLDEVLKHPMIGLAPPERNYLQRLSDQFREEDLDSGLFYLSDAN 189

Qy 283 SFVDAE-----WKVSDTRSGHFLRLLRWLQLSVQDTKLHLQNI 321

Qy 190 QFVDTTENRIMFPLTDQGKVIAFSGRIMQKTSQSKYKNSR---STAIKNKSYELYHM 246

Qy 322 HLHBRNRGKLAQYFAMMKDKKRKFKRRESLPEORSQMGAFDRD-AEYIRALRHLYVYP 380

Qy 247 DRAKRSRGKASEIYLMEGFMDVIAARYAGIENAVASMGTALESRBHEHLKRLTKLV---303

Qy 381 SARPVLSLSSGSWTKLHFFQEVLG---ASMRALYIVRDPRAWTYSMLYNKSPKSLYSLKHY 438

Qy 304 -----LYVDGDKGQAATLKAIDBIGDMNPVQTSMPDNLDPDEYLRQNG 347

Qy 439 PEHLAKLFKLEGGKCNLNSGTAFYBEPKRKELSKSKSNAVSLSHMLANTAA---494

Qy 348 PEDLAYLT---KTRISPIEFTHQYKPBENSENLQAQIEFLERIAPIVQESIAQNS 403

Qy 495 -LRTNDLILPTSYOLVKPFDIVH---PFQKTERIFAFGJPL- 533

Qy 404 YIHILADSLS-ASFDTYQLEQIVINESRQVYQRNRMEGISRPTPTMPVTKQLSAIMRAAEH 462

Qy 534 -----SPASLNQIL---FATSTNLFVLPYE---GEISP----TNTNWKQN 569

Qy 463 LLXRMMEPLVNDYLRLEDFAFATPEFQVLYDQYGNLPPVLAETOEEVERAWYQV 522

Qy 570 LPRD---EIKUIJEN 580

Qy 523 LAQDLPAEISPOELSEVENTRN 544

Search completed: June 23, 2005, 08:50:15
Job time : 16.5183 secs

RESULT 14
 T2168 hypothetical protein C53A5.9 - *Caenorhabditis elegans*
 C; Species: *Caenorhabditis elegans*
 C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C; Accession: T20168
 R; Mortimore, B.
 Submitted to the EMBL Data Library, November 1996
 A; Reference number: Z19332
 A; Accession: T20168
 A; Status: preliminary; translated from GB/EMBL/DDJB
 A; Molecule type: DNA
 A; Residues: 1-480 <WIL>
 A; Cross-references: UNIPROT:Q17700; EMBL:Z81486; PIDN:CA03989.1; GSPDB:GN000023; CESP:CS
 A; Experimental source: clone C53A5
 C; Genetics:
 A; Map Position: 5
 A; Introns: 28/1; 59/3; 194/1; 264/3; 332/1

Query Match 3.1%; Score 98; DB 2; Length 480;

Best Local Similarity 20.6%; Prod. No. 12; Mismatches 60; Indels 140; Gaps 22;

Qy 199 ELLDWWSCTCSOPI---CAKMRTRTEAGSKSKLSSEGHMIDLDPDVVITSUPGSGAEILKQL 255
Db 67 EVADAWDVQCPQLEYQDAGFSSESMETOKEL---EDCSSTEASRYFLENLQS-FKLMKIL 123Qy 256 ---PFNSSFDFLYRVPATAVIDPTELEIDSFVDAECKVSDTRSGHFLRPLLRLWMLQSIV 311
Db 124 QLTQSLNPQESSYNR---KIISTEMK-----KINDTVS-----QLKY 157Qy 312 QDTKLHLQNTLHL-----EPNRGKLA---QYFAMNK-----DKKRKEFKR 347
Db 158 AGKRIEMINHLEKRFTRPTGKDGALEPNDPVLIAQSAHKPLLISSFPKTQKTKP 217Qy 348 RESLPEQSRQSMQKAFDRDAFYIRALRRHLWYPPSARPVLSLSSGSWTKLKHFFQE-----402
Db 218 IGRDLPRAFNFAVASSKTNIV-----IGETHNGFLNKFEYNNQKQAR 262Qy 403 VLGASM-----ALTYRDPRAWIYSMLYNSKPSL-----KNUPE----440
Db 263 CMGANLNHKRTKPAAGPHNGALYVYGG-----YDSIVLSSVELYDLEEGTWNKPSLNNC 317Qy 441 -----HIAKLFKIEGGKGKCNL-----SSGAYFEYPLERKLXELSKSKSNAVSLSH 485
Db 318 RANAAVVACVGEIFTLGGFGKSNBESIEKISASNEFEGF-EMEGGSGSGFGCVCFQCR 376Qy 486 LWLA-----NTAAALRINTDLIPT---SYQLVKFEDIVHFQKTERIFAFGLPL 533
Db 377 IYIAGGWNSNTNTKSVRSYDPTYTTWDRPSMNQARKATFL---ATNEAIVARYGDB 433Qy 534 SPASLNOILPATSTNLFY--LPYEBISPPNTNWK 567
Db 434 ESALLDQIERDFPKLWKSSIVPSKPHVPPSYNTEK 469

RESULT 15

H95123 DNA primase [imported] - *Streptococcus pneumoniae* (strain TIGR4)C; Species: *Streptococcus pneumoniae*

C; Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C; Accession: H95123
 R; Tattelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidson, J.D.; Umayam, L.A.; White, O.; Saizberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, M.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001

THIS PAGE BLANK (USPTO)

| | | | | | | |
|--|----------------------------------|-------------------------|-------|--------|---------|---------------------|
| Scoring table: | BLOSUM62 | Gapop 10.0 , Gapext 0.5 | | | | |
| Searched: | 1612378 seqs, 512279187 residues | | | | | |
| Total number of hits satisfying chosen parameters: | 1612378 | | | | | |
| Minimum DB seq length: 0 | | | | | | |
| Maximum DB seq length: 2000000000 | | | | | | |
| Post-processing: | Minimum Match 0% | | | | | |
| | Maximum Match 100% | | | | | |
| | Listing first 45 summaries | | | | | |
| Database : | UniProt_03: | | | | | |
| | 1: uniprot_sprot; | | | | | |
| | 2: uniprot_trembl; * | | | | | |
| Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. | | | | | | |
| SUMMARIES | | | | | | |
| Result No. | Score | Query | Match | Length | DB ID | Description |
| 1 | 3142 | 99.8 | 1222 | 2 | Q8IZUB | Q8izub homo sapien |
| 2 | 1821 | 57.9 | 755 | 2 | Q9HAK5 | Q9hak5 homo sapien |
| 3 | 920 | 29.2 | 480 | 2 | Q9H51 | Q9h51 homo sapien |
| 4 | 863 | 27.4 | 784 | 2 | Q6PS23 | Q6ps23 homo sapien |
| 5 | 615 | 19.5 | 314 | 2 | Q9HAD7 | Q9had7 homo sapien |
| 6 | 150 | 4.8 | 411 | 2 | Q9QCO | Q9qco homo sapien |
| 7 | 147 | 4.7 | 411 | 2 | Q43916 | Q43916 homo sapien |
| 8 | 136 | 4.3 | 958 | 2 | Q9UL01 | Q9ul01 homo sapien |
| 9 | 136 | 4.3 | 958 | 2 | Q9BL14 | Q9bl14 homo sapien |
| 10 | 134 | 4.3 | 483 | 2 | Q794G9 | Q794g9 mus musculus |
| 11 | 134 | 4.3 | 530 | 2 | Q882T6 | Q882t6 mus musculus |
| 12 | 134 | 4.3 | 530 | 2 | Q9WV3 | Q9wv3 mus musculus |
| 13 | 133 | 4.2 | 483 | 2 | Q9ED5 | Q9ed5 homo sapien |
| 14 | 133 | 4.2 | 530 | 2 | Q9Y4C5 | Q9y4c5 homo sapien |
| 15 | 129 | 4.1 | 304 | 2 | Q677T4 | Q677t4 lymphocyte |
| 16 | 128 | 4.1 | 486 | 2 | Q9VMC3 | Q9vmc3 drosophila |
| 17 | 127 | 4.0 | 441 | 2 | Q934O3 | Q934o3 torpedo cat |
| 18 | 126.5 | 4.0 | 420 | 2 | Q9DBy9 | Q9db9 brachydanio |
| 19 | 126 | 4.0 | 388 | 2 | Q9WUE5 | Q9wue5 mus musculus |
| 20 | 125 | 4.0 | 388 | 2 | Q9R1I1 | Q9r1i1 mus musculus |
| 21 | 123.5 | 3.9 | 395 | 2 | Q9GZK3 | Q9gzk3 homo sapien |
| 22 | 122.5 | 3.9 | 384 | 2 | Q62E8B8 | Q62e8b burholderi |
| 23 | 122.5 | 3.9 | 479 | 2 | Q7IGC8 | Q7igc8 homo sapien |
| 24 | 122 | 3.9 | 392 | 2 | Q9GN39 | Q9gn3 xenopus lae |
| 25 | 122 | 3.9 | 515 | 2 | Q9BV31 | Q9bv31 renealmia s |
| 26 | 121.5 | 3.9 | 392 | 2 | Q63I21 | Q63i21 burholderi |
| 27 | 121 | 3.8 | 520 | 2 | Q672W9 | Q672w9 curculigo c |
| 28 | 120.5 | 3.8 | 479 | 2 | Q75099 | Q75099 homo sapien |
| 29 | 119.5 | 3.8 | 390 | 2 | Q7LCN3 | Q7lcn3 homo sapien |
| 30 | 119.5 | 3.8 | 411 | 2 | Q8GS9 | Q8gs9 homo sapien |
| 31 | 119.5 | 3.8 | 518 | 2 | Q9GHE3 | Q9ghe3 drymophila |
| 32 | 119 | 3.8 | 558 | 2 | Q88199 | Q88199 mus musculus |
| 33 | 119 | 3.8 | 377 | 2 | Q79415 | Q79415 mus musculus |
| 34 | 118.5 | 3.8 | 395 | 2 | Q7NRP85 | Q7nrp85 chromobacte |
| 35 | 118 | 3.7 | 395 | 2 | Q9QUP4 | Q9qup4 mus musculus |
| 36 | 118 | 3.7 | 515 | 2 | Q6LA24 | Q6la24 curculigo c |
| 37 | 117 | 3.7 | 515 | 2 | Q8HV32 | Q8hv32 renealmia t |
| 38 | 117 | 3.7 | 559 | 2 | Q9FGV3 | Q9fgv3 arabidopsis |
| 39 | 116.5 | 3.7 | 530 | 2 | Q9DOK5 | Q9dok5 mus musculus |
| 40 | 116 | 3.7 | 363 | 2 | Q9VN4C | Q9vn4c droophila |
| 41 | 116 | 3.7 | 515 | 2 | Q8HV29 | Q8hv29 renealmia t |
| 42 | 116 | 3.7 | 515 | 2 | Q8HV34 | Q8hv34 etingeria e |
| 43 | 115 | 3.7 | 515 | 2 | Q8HV65 | Q8hv65 etingeria e |
| 44 | 114 | 3.6 | 515 | 2 | Q8HV21 | Q8hv21 sili-quamomu |
| 45 | 114 | 3.6 | 515 | 2 | Q8HV30 | Q8hv30 renealmia s |

| | | | |
|--|---|--|--------------|
| Db | 308 | MNRYNGAMMDWDAHYKMFWDHGNSPMASIQEQAABFKRWTQFVNVTQMESTIT | 367 |
| Qy | 61 | RIAVYFGPYINSSCRFIDSSNPGLOISLAVNNTTEHVYSIVTDYHNLTKRNFYLGFF | 120 |
| Db | 368 | RIAVYFGPYINSSCRFIDSSNPGLOISLAVNNTTEHVYSIVTDYHNLTKRNFYLGFF | 427 |
| Qy | 121 | ASVADQGQTTRFLGLGTQAIVKPVHDR1IIPFGKFENIAVLTF 173 | |
| Db | 428 | ASVADQGQTTRFLGLGTQAIVKPVHDR1IIPFGKFENIAVLTF 480 | |
| RESULT 4 | | | |
| Q6PZ3 | Q6PZ3 | PRELIMINARY; PRT; 784 AA. | |
| AC | Q6PZ3; | SEQUENCE FROM N.A. | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Created) | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Last sequence update) | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Last annotation update) | |
| DE | C18orf4 Protein (Fragment). | | |
| GN | Name=C18orf4; | | |
| OS | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| NCBI_TaxID | 9606; | | |
| RN | [1] | | |
| RP | | | |
| RC | | | |
| TISSUE=Breast; | | | |
| RX | PMID=1238857; PubMed=12477932; DOI=10.1073/pnas.242603899; | | |
| RA | Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., | | |
| RA | Klaasner R.D., Collin F.S., Wagner L., Shemesh C.M., Schuller G.D., | | |
| RA | Altachui S.F., Zeeberg B.R., Butow K.H., Schatz N.K., | | |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., | | |
| RA | Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., | | |
| RA | Stapleton M., Soares M.F., Bonaldo M.F., Casavant T.L., Scheetz T.E., | | |
| RA | Brownstein M.J., Usdin T.B., Yoshiyuki S., Carninci P., Prange C., | | |
| RA | Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallahy S.J., | | |
| RA | Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., | | |
| RA | Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., | | |
| RA | Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., | | |
| RA | Whiting M., Heilcon E., Ketteman M., Madan A., Rodrigues S., Sanchez A., | | |
| RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | | |
| RA | Rodriguez P.J., Grimes J., Schmutz J., Butterfield Y.S., | | |
| RA | Krzewinski M.I., Marra M.A.; Jones S.J., Marra M.A.; Schein J.E., | | |
| RA | Jones S.J., Marra M.A.; Generation and initial analysis of more than 15,000 full-length human | | |
| RT | and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). | | |
| RL | | | |
| RN | [2] | | |
| RP | | | |
| RC | | | |
| TISSUE=Breast; | | | |
| RA | Strausberg R.; Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases. | | |
| DR | EMBL; BC062557; AA062557.1; -. | | |
| FT | NON_TER | SEQUENCE 784 AA: 89666 MW: 2A86FFC4DDCB846C CRC64; | |
| FQ | 27.4%; Score 863; DB 2; Length 784; | | |
| Best Local Similarity 99.4%; Pred. No. 3.e-60; | | | |
| Matches 161; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | | | |
| Qy | 1 | MNRYNGAMMDWDAHYKMFWDHGNSPMASTOEAOAAEFKRTQFVNVTQMESTIT | 60 |
| Db | 617 | MNRYNGAMMDWDAHYKMFWDHGNSPMASTOEAOAAEFKRTQFVNVTQMESTIT | 676 |
| Qy | 61 | RIAYVFGPYINSSCRFIDSSNPGLOISLAVNNTTEHVYSIVTDYHNLTKRNFYLGFF | 120 |
| Db | 677 | RIAYVFGPYINSSCRFIDSSNPGLOISLAVNNTTEHVYSIVTDYHNLTKRNFYLGFF | 736 |
| Qy | 121 | ASVADQGQTTRFLGLGTQAIVKPVHDR1IIPFGKFENIAVGL 162 | |
| Db | 737 | ASVADQGQTTRFLGLGTQAIVKPVHDR1IIPFGKFENIAVGL 778 | |
| RESULT 5 | | | |
| Q9HAD7 | Q9HAD7 | PRELIMINARY; | PRT; 314 AA. |
| AC | Q9HAD7; | | |
| DT | 01-MAR-2001 | (TREMBLrel. 16, Created) | |
| DT | 01-OCT-2001 | (TREMBLrel. 16, Last sequence update) | |
| DT | 01-OCT-2001 | (TREMBLrel. 18, Last annotation update) | |
| DE | Hypothetical protein FLJ11787. | | |
| OS | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| NCBI_TaxID | 9606; | | |
| RN | [1] | | |
| RP | | | |
| RC | | | |
| TISSUE=Whole embryo; | | | |
| RX | PubMed=14703039; DOI=10.1038/ng1285; | | |
| RA | Ota T., Suzuki Y., Niishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J., Saito K., Kawai Y., Nakamura Y., Nagahashi K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M., Shiratori A., Sudou H., Hobo T., Kaku H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M., Yamazaki M., Niromiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kitamata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotaka T., Kusano J., Kanehori K., Takakashi-Fujii A., Hara H., Tanase T., Nomura Y., Togoya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Oshima A., Sasaki N., Aotsuka S., Yoshihara Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami T., Terashima Y., Suzuki O., Nakagawa S., Seno A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Kumagai A., Takekoto M., Kawakami B., Yamaizaki M., Watanabe K., Kumagai A., Itakura S., Fujizumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikili T., Kobatake N., Inagaki H., Ikeda Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Konatsu T., Mizushima-Sugano J., Sato T., Shirai Y., Nakagawa K., Okumura K., Nagase T., Nomura K., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S., "Complete sequencing and characterization of 21,243 full-length human cDNAs"; Nat. Genet. 36:40-45 (2004); RL Genet. 36:40-45 (2004); DR EMBL; AK021849; BAB13912.1; -. FT NON_TER 314 MW: 722B5555DC05538 CRC64; | | |
| SQ | SEQUENCE 314 AA: | | |
| Query Match | 19.5%; Score 615; DB 2; | | |
| Best Local Similarity 100.0%; Pred. No. 7.6e-41; | | | |
| Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| Qy | 224 | KKSISSEGHMMDPDVITSLPGSGAEIKLQKPFNSDFLYIRTPAYDIPETLEIDS 283 | |
| Db | 198 | KKSISSEGHMMDPDVITSLPGSGAEIKLQKPFNSDFLYIRTPAYDIPETLEIDS 257 | |
| Qy | 284 | FVDACEWKSDIRSFGFRLLRGMLQSLYQDTKLHLQNIHLHEPNRGKLAQYFANMKD 340 | |
| Db | 258 | FVDACEWKSDIRSFGFRLLRGMLQSLYQDTKLHLQNIHLHEPNRGKLAQYFANMKD 314 | |
| RESULT 6 | | | |
| Q9EQC0 | Q9EQC0 | PRELIMINARY; | PRT; 411 AA. |
| AC | Q9EQC0; | | |
| DT | 01-MAR-2001 | (TREMBLrel. 16, Created) | |
| DT | 01-MAR-2001 | (TREMBLrel. 16, Last sequence update) | |
| DT | 25-OCT-2004 | (TREMBLrel. 28, Last annotation update) | |

| | |
|---|---|
| DE Galactose 6-O-sulfotransferase GST-1 (Carbohydrate sulfate transferase 1). | Db 124 LRSLYDDLYFLENYIKPPPVNHTTNRVFRGASRVICSRVCDPSSDILLEGDCVR 183 |
| DN Name=Cat1; Synonyms=Cat1; | Qy 396 KLIHFFQEVIGA-----SMRALIVDRPRAVYISM 424 |
| OS Mus musculus (Mouse) | Db 184 MCCLLNITLAAACRERESHVAIKTVRYPEVNDLRVALYEDPRINLKVNQLVRDERTIASR 243 |
| OC Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus. | Qy 425 LYNSKPLSISLNKVNPEHLAKLPIEGKGKCNLNSGYAFEPRLRKLSK-----SK 476 |
| OX NCBI_TaxID=10090; | Qy [1] |
| RN [1] | RP SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Mammary gland; MEDLINE=200602097; PubMed=11093/glycob/11.1.75; |
| RC RA Hemmerich S.; Lee J.-K.; Bhakta S.; Bistrup A.; Ruddle N.R., Rosen S.D.; | Db 244 SETFRY-----RWRWLYCTGR-----KPYNUJYOTLJTVCFESS 281 |
| RA "Chromosomal localization and genomic organization of the galactose/N-acetylgalactosamine/N-acetylglucosamine 6-O sulfotransferase gene family." Glycobiology 11:75-87 (2001). | Qy 477 SNAVSLIHLWLNATAALRINTDLPLPTSYOLVKFEDIVHFPQKTRIFAFGIP----- 533 |
| RT RL | Db 282 SVSTGLRPSWLKG-----KYMVLVRYELARNPMKKTETEYEEFGIPDSH 327 |
| RT RN | Qy 534 -----SPASLNCILPATSTNLFLYLPYEGETISPNTNTWQKQNPDEIKLIEJCW 583 |
| RT RN | Db 328 VAHWIQNNTRGDPFLGHKYSTVRN-----SAATAEKMRFRILSYDVTAVAQNACQ 377 |
| RN [2] | RP SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Mammary gland; MEDLINE=200602097; PubMed=12477932; DOI=10.1073/pnas.242603899; |
| RC RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collinge F.S., Wagner L., Schenner C.M., Schulter G.D., Altschuler S.F., Zeeberg B., Buetow K.H., Schaeffer C.P., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stappert M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullally S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.X., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Paley J., Helton R., Kerteszen M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.N., Touchman J.W., Green G.E., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krywinski M.I., Skalska U., Smailus D.E., Schnierch A., Schein J.B., Jones S.J., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). | |
| RT RN [3] | RP SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Mammary gland; MEDLINE=200202002; to the EMBL/GenBank/DDBJ databases. |
| RC RA Strausberg R.; Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases. | Db 043916 PRELIMINARY; PRT; 411 AA. |
| RL DR EMBL; AP280087; AAG4245.RP; | DT 01-JUN-1998 (TRIMBLrel. 06, Created) |
| DR EMBL; BC030067; AAH300667.1. -; | DT 01-JUN-1998 (TRIMBLrel. 06, Last sequence update) |
| DR MGFI; MGI:1924219; Chst1. | DT 25-OCT-2004 (TRIMBLrel. 28, Last annotation update) |
| DR GO:0005615; c:extracellular space; TAS. | DE Chondroitin 6-sulfotransf erase (Keratan sulfate Gal-6-sulfotransf erase) (Carbohydrate sulfotransf erase 1) (Carbohydrate sulfotransf erase 1) |
| DR InterPro:IPR000863; Sulfotransferase_1. | DE (Keratan sulfate Gal-6) sulfotransf erase 1. |
| DR PFam; PF00685; Sulfotransferase_1. | DE Name=CHST1; |
| SQ SEQUENCE 411 AA; 46903 MW; B1AE590EF5B9CBDC CRC64; | GN Homo sapiens (Human) |
| RT RT | OS Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominidae; Homo. |
| RT RN [1] | OC TAXID=9606; |
| RT RN | RN SEQUENCE FROM N.A. |
| RT RN | RP MEDLINE=98-016085; PubMed=9639683; DOI=10.1016/S0925-4439(98)00028-3; |
| RT RN | RA Mazany K.D., Peng T., Watson C.E., Tabas I., Williams K.J.; RT "Human chondroitin 6-sulfotransf erase: Cloning, gene structure, and chromosomal localization.", |
| RT RN | RL Biochim. Biophys. Acta 1407: 92-97 (1998). |
| RT RN [2] | RN SEQUENCE FROM N.A. |
| RT RN | RP MEDLINE=971189336; PubMed=9117134; |
| RT RN | RA Williams K.J.; RT "Atherosclerosis: cell biology and lipoproteins.", |
| RT RN | RL Curr. Opin. Lipidol. 7:U202-U208 (1996). |
| RT RN [3] | RN SEQUENCE FROM N.A. |
| RT RN | RC TISSUE:Brain; |
| RT RN | RP MEDLINE=98070405; PubMed=9405439; DOI=10.1074/jbc.272.51.32321; |
| RT RN | RA Fukuta M., Inazawa J., Tsuzuki K., Shimada E., Habuchi O.; RT "Molecular cloning and characterization of human keratan sulfate Gal-6-sulfotransf erase.", |
| RT RN | RL J. Biol. Chem. 272:32321-32328 (1997). |
| RT RN [4] | RP SEQUENCE FROM N.A. |
| RT RN | RC TISSUE:Umbilical vein endothelium; |
| RT RN | RP MEDLINE=99168906; PubMed=1049591; DOI=10.1006/geno.1998.5653; |
| RT RN | RA Li X., Tedder T.F.; RT "CHST1 and CHST6 sulfotransf erases expression by human vascular endothelial cells: cDNA cloning, expression, and chromosomal localization.", |
| RT RN | RL Genomics 55:345-347 (1999). |
| RT RN [5] | RP SEQUENCE FROM N.A. |
| RT RN | RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W., Robertson P.D., Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Sherwood J.K., Wittek L.A., Nickerson D.A.; |
| RT RN | RA Schackwitz W.S., Sherwood J.K., Wittek L.A., Nickerson D.A.; |

Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.

| | | | |
|---------------|--|------------|--|
| RL | 417 PRAWIYSMLYNKPSLSSLYKVNPEHLAKLFKEGGKGKCNLNQYAFFYEPLRKELLSKSK 476 | Qy | 417 PRAWIYSMLYNKPSLSSLYKVNPEHLAKLFKEGGKGKCNLNQYAFFYEPLRKELLSKSK 476 |
| RN | [6] | Db | 236 PRGLASSETFDTY-----RLWRLVYGTGR----KPYNLDTQL-TTCBDF 219 |
| RC | SEQUENCE FROM N.A. | | |
| TISSUE=Brain; | | | |
| RX | MEDLINE=2388557; PubMed=12477932; DOI=10.1073/pnas.242603899; | Qy | 477 SNAVS--LISHLMANTAAALRINTDILPPTSYOLVKFEDIVHPQKTTTERIFAFGLIPU- 533 |
| RA | Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., | Db | 280 SNSVSTGMRPPMKKG-----KMLVRLVRDLARNPMKCTEEIVGFLGPID 325 |
| RA | Klaunser R.D., Collings F.S., Wagner L., Shenmam C.M., Schulter G.D., | Qy | 534 -----SPASINQILPATSTNLFYLVGEISPTNTNVKQNLPRDEKLHENI 581 |
| RA | Altenschul S.F., Zeeberg B.R., Buettow M.F., Schaefer C.F., Bhat N.K., | Db | 326 SHYARWIQNNTRDPTLGKHKYGTVRN-----SAATAEKWRFLSYDIVAFAGNA 375 |
| RA | Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., | Qy | 582 CWTLMDRIGY 591 |
| RA | Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., | Db | 376 CQQVLAQGY 385 |
| RA | Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., | Qy | |
| RA | Blakesley R.W., Uedin T.B., Toshiyuki S., Carninci P., Prange C., | Db | |
| RA | Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahay S.J., | Qy | |
| RA | Bobak S.A., McEwan P.J., McKernan K.J., Malek J., Gunaratne P.H., | Db | |
| RA | Richards S.W., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., | Qy | |
| RA | Villalon D.K., Muzyz D.M., Sodergren E.J., Lu X., Gibbs R.A., | Db | |
| RA | Fahy J., Heitton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., | Q9U01 | PRELIMINARY; |
| RA | Whiting M., Madan A., Young A.C., Bouffard G.G., | AC | Q9U01 |
| RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | AC | Q9U01 |
| RA | Rodriguez A.C., Grimes J., Schmitz J., Myers R.M., Butterfield Y.S., | AC | Q9U01 |
| RA | Krzewinski M.I., Skalska U., Smalius D.E., Schnarch A., Schein J.E., | AC | Q9U01 |
| RA | Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human | DT | 01-MAY-2000 (TREMBrel. 13, Created) |
| RT | and mouse cDNA sequences"; | DT | 01-MAY-2000 (TREMBrel. 13, Last annotation update) |
| RT | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). | DT | 05-JUL-2004 (TREMBrel. 27, Last annotation update) |
| RL | | DE | Squamous cell carcinoma antigen recognized by T cell (SART2 protein). |
| RN | | DE | Name=SART-2; |
| RP | | GN | Homo sapiens (Human). |
| RP | | OS | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| RP | | OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| RN | | NCBT_TaxID | 9606; |
| RN | | RN | [1] |
| RP | SEQUENCE FROM N.A. | RP | SEQUENCE FROM N.A. |
| RC | TISSUE=Brain; | RX | SEQUENCE FROM N.A. |
| RA | Strausberg R./ Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases. | RA | MEDLINE=20143850; PubMed=10679095; |
| RL | Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases. | RA | Nakao M., Shichijo S., Imaizumi T., Inoue Y., Matsunaga K., Yamada A., |
| RN | [8] | RA | Kikuchi M., Tsuda N., Ohta K., Takamori S., Yamana H., Fujita H., |
| RP | SEQUENCE FROM N.A. | RA | Itoh K. |
| RC | TISSUE=Brain; | RA | "Identification of a gene coding for a new squamous cell carcinoma antigen recognized by the CTL."; |
| RA | Strausberg R./ Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases. | RA | RL |
| DR | GO:0016021; C: integral to membrane; TAS. | RP | SEQUENCE FROM N.A. |
| DR | GO:0008146; P: sulfotransferase activity; TAS. | RX | TISSUE="testis"; MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899; |
| DR | InterPro:IPR000867; P: polysaccharide metabolism; TAS. | RA | Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collings F.S., Wagner L., Shenmam C.M., Schulter G.D., Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heleb F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson P.H., Mullahay S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Maruyama D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Heitton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krywinski M.I., Skalska U., Smalius D.E., Schnarch A., Schein J.E., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; |
| RL | RT | RA | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). |
| RN | | RN | [3] |
| RP | SEQUENCE FROM N.A. | RP | SEQUENCE FROM N.A. |
| RC | TISSUE=Brain; | RC | SEQUENCE FROM N.A. |
| RA | Strausberg R./ Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases. | RA | RT |
| DR | EMBL: AF098056; AA000871; -; | RA | RT |
| DR | EMBL: BC039245; AAH39245; -; | RA | RT |
| DR | Genew: HGNC:21144; SART2; | RA | RT |
| GO | GO:0005783; C: endoplasmic reticulum; TAS. | DR | RT |

| | | | | | |
|-----------------------|--|---|-----------------------------|-------------------------|-------------|
| Qy | SEQUENCE | 958 AA; | 109772 MW; | A3D05C6194B4D2F9 CRC64; | |
| Query Match | | 4.3%; | Score 136; | DB 2; | Length 958; |
| Best Local Similarity | | 31.8%; | Pred. No. 0.11; | | |
| Matches | 42; | Conservative | 17; | Mismatches | 51; |
| | | | | Indels | 22; |
| | | | | Gaps | 4; |
| Qy | 5 NGAMMDVDAHYKMFWDHGNSPMASIQEAEQAFKKRWTQVNVTFOQUESTIRAY | 64 | | | |
| Db | 603 HGAFTIQRDGLKYMWMDTGYSEKATPASITVPRGPNYVNTMHRSPITRAAY | 662 | | | |
| Qy | 65 VFYGPYINVSSCRFDSSNPGLQISLNVNNTTEHVSVTVDHNLTFRNYL-----GFG | 118 | | | |
| Db | 663 LFTGPSIDQS-----PTVHGDSQQLDVPIATSKRAYAT---YLMTGEATGOS | 707 | | | |
| Qy | 119 GFSV-ADQGQI | 129 | | | |
| Db | 708 AFAGVIADEHKI | 719 | | | |
| RESULT 9 | | | | | |
| QBLI4 | PRELIMINARY; | PRT; | 958 AA. | | |
| AC | QBLI4; | | | | |
| DT | 01-MAR-2003 | (TREMBLe1. | 23; Created) | | |
| DT | 01-MAR-2003 | TREMBLe1. | 23; Last sequence update) | | |
| DT | 01-MAR-2004 | (TREMBLe1. | 28; Last annotation update) | | |
| DE | Mus musculus | 9.5 days embryo parthenogenote cDNA, enriched library, clone:B13024B19 product:SQAMOUS CELL CARCINOMA ANTIGEN RECOGNIZED BY T CELL homolog (Squamous cell carcinoma antigen recognized by T cells 2). | | | |
| DE | Name Sar2; | | | | |
| GN | OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| OC | QX | QX-TAXID=10090; | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=C57BL/6J; TISSUE=Parthenogenote; | | | | |
| RC | MEDLINE:99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9; | | | | |
| RA | Carninci P., Hayashizaki Y. | | | | |
| RT | "High-efficiency full-length cDNA cloning."; | | | | |
| RL | Meth. Enzymol. 303:19-44 (1999). | | | | |
| RN | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=C57BL/6J; TISSUE=Parthenogenote; | | | | |
| RC | MEDLINE:21085660; PubMed=11217851; DOI=10.1038/35055500; | | | | |
| RA | "Functional annotation of a full-length mouse cDNA collection." | | | | |
| RT | Nature 409:685-690 (2001). | | | | |
| RN | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=C57BL/6J; TISSUE=Parthenogenote; | | | | |
| RC | MEDLINE:20499374; PubMed=11042159; DOI=10.1101/gr.145100; | | | | |
| RA | Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y. | | | | |
| RT | "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes." | | | | |
| RL | Genome Res. 10:1617-1630 (2000). | | | | |
| RN | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=C57BL/6J; TISSUE=Parthenogenote; | | | | |
| RC | MEDLINE:20530913; PubMed=11076861; DOI=10.1101/gr.152600; | | | | |
| RA | Shibata K., Itoh M., Aizawa K., Nagaoka S., Sakai N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., | | | | |
| Qy | 5 NGAMMDVDAHYKMFWDHGNSPMASIQEAEQAFKKRWTQVNVTFOQUESTIRAY | 64 | | | |
| Db | 603 HGALIRQDRGLYKMYNMDDTGSEKANFASTMYPRGYPYNGTNNVNTMHLRSPTTAY | 662 | | | |
| Qy | 65 VFYGPYINVSSCRFDSSNPGLQISLNVNNTTHVVSVTDTHNLKTRFNYL-----GFG | 118 | | | |
| Db | 663 LFQPSVVDQS-----FSIHDGPQLDVFTATSEHAYAT--YLMTGENTGHS | 707 | | | |
| Qy | 119 GRASV-ADQGQI | 129 | | | |
| Db | 708 AFQVIADEHKI | 719 | | | |

| | | |
|--|--|---|
| Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimaldi J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.J., Smailus D.B., Schnech A., Schein J.E., Jones S.J., Marrs M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [2] | SEQUENCE FROM N.A. STRAIN=C57BL/6; TISSUE=Brain; Strauberg R.; Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases. | RT acetylglucosamine at the nonreducing end of an N-acetyllactosamine sequence."; RT Biochim. Biophys. Acta 1523:269-276 (2000). RT EMBL; AB014679; BAA34255.2; DR EMBL; AB021124; BAB16866.1; DR EMBL; AB021125; BAB16887.1; DR GO; GO:0008146; F:sulfotransferase activity; IEA. DR GO; GO:0016740; F:transf erase activity; IEA. DR InterPro; IPR00863; Sulfotransferase. DR Pfam; PF00685; Sulfotransfer-1. 1. KW Transf erase. SQ SEQUENCE 483 AA; 52787 MW; E27797D4931BA18 CRC64; |
| | | Query Match Score 4.2%; DB 2; Length 483; Best Local Similarity 26.2%; Pred. No. 0.071; Gaps 9; Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9; |
| | | Qy 408 MRALYIVDPRAVTNSLYNSKPSLY-SLKVNVP-----HLAKLFKEGGK-GKCNLNS 459 Db 278 LKVTHLVRDPPA-VASSRIRSRLGLIRESLQVRSRDPRAHMPFLEAGHKLGAKEGV 336 |
| | | Qy 460 GYAFFEYEPRLKELSKSNSNAVSLSHWLANTAAALINTDLPLTSYOLVKFEDIVHFPQ 519 Db 337 GGPDADYHAL-----GAMEVICNSMALKTLQTA-----VSARNATQAANAWRTALT 441 |
| | | Qy 520 KTERIFAAFLGIPILSPA---SLNQILPATSTNLFLYPEGEISPTN---TNWKQNLP 571 Db 387 KTLLRVIDPGVLVSPMEQFALNMTSGSSSKPFV----VSARNATQAANAWRTALT 441 |
| | | Qy 572 RDETKLIEENICWTLMRDLRGYPK 593 Db 442 FQQIKQVEEFCYQPMAVLGYER 463 |
| | | RESULT 14 Q9Y4C5 PRELIMINARY; PRT; 530 AA. AC Q9Y4C5; Q9GZNS; Q9Y6F2; DT 01-NOV-1999 (TREMBrel. 12, Created) DT 01-MAR-2003 (TREMBrel. 23, Last sequence update) DT 05-JUL-2004 (TREMBrel. 27, Last annotation update) DE Long form of N-acetylglucosamine-6-O-sulfotransferase (GlcNAcST) DE (Carbohydrate sulfotransferase 2). DE Name=GNEST; Synonyms=CHSTR2; DN Q9Y4C5 GN Homo sapiens (Human). OS Homo sapiens (Human). OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCBI_TaxID=9606; RN [1]- |
| | | RP TISSUE=brain; RN MEDLINE:93391845; PubMed=9722682; RN Sequence FROM N.A. RN MEDLINE:991658906; PubMed=10049591; DOI=10.1006/geno.1998.5653; RN Uchiyama K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q., RN Ochiai T., Nakagawa A., Kadomatsu K., Habuchi O., Yokoyama I., Yamamura K., RN Mitsuoka C., Kannagi R., Habuchi O., Yokoyama I., Yamamura K., RN Ozaki T., Nakagawa A., Kadomatsu K., Muramatsu T., RN Li X., Tedder T.F.; "Human N-acetylglucosamine-6-O-sulfotransferase involved in the RN biosynthesis of 6-sulfo sialyl Lewis X: molecular cloning, chromosomal RN mapping, and expression in various organs and tumor cells.", RN J. Biochem. 124:670-678(1998). RN [2] RN Sequence FROM N.A. RN Genomics 55:145-147(1999); RN EMBL; AB014680; BAA34266.2; RN EMBL; AF083066; ADD20981.1; RN Genew; HGNC:1970; CHSTR. RN GO; GO:0008146; F:transf erase activity; IEA. RN GO; GO:0016740; F:transf erase activity; IEA. |
| | | RESULT 13 D Q9UEDS PRELIMINARY; PRT; 483 AA. C Q9UEDS; 01-MAY-2000 (TREMBrel. 13, Created) T 01-MAR-2003 (TREMBrel. 23, Last sequence update) T 05-JUL-2004 (TREMBrel. 27, Last annotation update) T Name=GN6ST; S Homo sapiens (Human). C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. C NCBI_TaxID=9606; X Uchiyama K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q., X Mitsuoka C., Kannagi R., Habuchi O., Yokoyama I., Yamamura K., X Ozaki T., Nakagawa A., Kadomatsu K., Muramatsu T., X "Human N-acetylglucosamine-6-O-sulfotransferase involved in the X biosynthesis of 6-sulfo sialyl Lewis X: molecular cloning, chromosomal X mapping, and expression in various organs and tumor cells.", X J. Biochem. 124:670-678(1998). X Sequence FROM N.A. X TISSUE=Placenta; X MEDLINE=20498786; PubMed=11042394; DOI=10.1016/S0304-4165(00)00136-7; X Sakaguchi H., Kitagawa H., Sugahara K.; "Functional expression and genomic structure of human N- X acetylglucosamine-6-O-sulfotransferase that transfers sulfate to b-N- |

DR InterPro; IPRO000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransferase_1.
 KW Transferase.
 SQ SEQUENCE 530 AA; A82CA22TB9D5651B CRC64;
 Query Match Score 133; DB 2; Length 530;
 Best Local Similarity 26.2%; Pred. No. 0.08; Mismatches 36; Indels 81; Gaps 9;
 Matches 53; Conservative 36; N mismatches 81; Indels 32; Gaps 9;
 Qy 408 MRALYIVRPRAWIYSMLYNSKPSLY-SLKNYPE-----HLAKLFKTEGGK-GKCNLNS 459
 Db 325 LKVTHLVRDRA-VASSRIRSRRGIRESLQVWRSRDPRAHMPFLEAGHKLGAKCEGV 383
 Qy 460 GYAFYEYEPRLKELSKSKSNSAVSLSHLWANTAAALRINTDLPLTSYQLVKFEDIVHFPQ 519
 Db 384 GGPDYHAL-----GAEVICNSMAKTLQTLAQ-PPDWLQGHYLIVRVYEDLVGPV 433
 Qy 520 KTERIFAFGIPILSPA---SINQILPATSTNLFLYLPYEGEISPTN---TNNWKQNLP 571
 Db 434 KTLRKYDFGLLSPMEPMOFALNMNTSGSSSSKRPV----VSARNATQAANAWRTALT 488
 Qy 572 RDEIKLIENTCWTLMDRIGYPK 593
 Db 489 FQQKIQKVEEEFCYQPMHAVLGTER 510

RESULT 15

| | | | |
|---|--------------|------|-------------------------|
| Q6777T4 | PRELIMINARY; | PRT; | 304 AA. |
| ID Q6777T4; | | | |
| AC Q6777T4; | | | |
| DT 25-OCT-2004 | (TREMBrel. | 28, | Created) |
| DT 25-OCT-2004 | (TREMBrel. | 28, | Last sequence update) |
| DT 25-OCT-2004 | (TREMBrel. | 28, | Last annotation update) |
| DE Hypothetical protein. | | | |
| OS Lymphocytis disease virus - isolate China. | | | |
| OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Lymphocystivirus. | | | |
| OX NCBI_TaxID:256729; | | | |
| RN [1] | | | |
| RP SEQUENCE FROM N.A. | | | |
| RX PubMed=15194775; DOI=10.1128/JVI.78.13.6982-6994.2004; | | | |
| RA Zhang Q.Y., Xiao P., Xie J., Li Z.Q., Gui J.F.; | | | |
| RT "Complete Genome Sequence of Lymphocytis Disease Virus Isolated from China." | | | |
| RT J. Virol. 78:6982-6994 (2004). | | | |
| RL EMBL; AY80826; AAI11023.1; -. | | | |
| DR InterPro; IPRO000863; Sulfotransferase. | | | |
| DR Pfam; PF00685; Sulfotransferase_1. | | | |
| KW Hypothetical protein. | | | |
| SQ SEQUENCE 304 AA; 35898 MW; 63C9BDDBFB2145C22 CRC64; | | | |

Query Match Score 1.1%; DB 2; Length 304;
 Best Local Similarity 21.1%; Pred. No. 0.077;
 Matches 86; Conservative 62; N mismatches 126; Indels 134; Gaps 22;

| | | | |
|--|--|--|--|
| Qy 188 ILLVITALWPIELDWSTCSQICAKWTRTEAEGSKKSISSEHHMDLUPDVVITSLPGS 247 | | | |
| Db 8 MLLIILLLML---IW-IFQRPV-----YECKPNK-----ILLITTRS 43 | | | |
| Qy 248 GAEILKOLPNSSPDLYIRVPTAYIDIPETBLEIDSFDVDAEWKVDIIRRSHFRLLRGWL 307 | | | |
| Db 44 GSSFGEIIFSRSRSDVFLFPL-----WHLINEIVTEYVKVLKF 83 | | | |
| Qy 308 Q-SLYQDTKLHLQNTH-----LHEPNRKGLAQYFAMNDKCRKFERESLP---EQR 355 | | | |
| Db 84 NCELVVLRLKTQNFFKRMSYSKALCKP-GKTCY-GLNGDRKTYCDGKCPQLNLDA 140 | | | |
| Qy 356 SQMKGAFDRDAEYIALRRLVVYPSARPVLSLSSGSWTKLHFQEVLGASMBALEYVR 415 | | | |
| Db 141 SLYCOTF-DTVVVKTVR-----IRAKTOAEL---MTOFDKIHLYR 179 | | | |
| Qy 416 DPRAWIYSMLYNSKPSLYSISKNVPBHLAKUFKIEGKGKCNLNSSYAFEPRLKELSKS 475 | | | |
| Db 180 DPRG-----SFNSK-----IKTF-----NRDYNF-----KQIAKI 204 | | | |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 08:32:43 ; Search time 42.6885 Seconds

(without alignments)
3578.727 Million cell updates/sec

Title: US-10-697-828-13

Perfect score: 2056

Sequence: 1 MWLPRVSSTAVTLIAQTF.....LPRGLNGFTWASSTASHPRN 395

Scoring table: BLOSUM62

Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqP1980s:*
- 2: geneseqP1990s:*
- 3: geneseqP2000s:*
- 4: geneseqP2001s:*
- 5: geneseqP2002s:*
- 6: geneseqP2003ab:*
- 7: geneseqP2003bs:*
- 8: geneseqP2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description | | | | | |
|------------|-------|-------|-------|--------|----------|----|--------------------|----|---------------|----|-------------------------------|--|
| 1 | 2056 | 100.0 | 395 | 4 | AYT2640 | | Aay72640 Human gly | OS | Homo sapiens. | | | |
| 2 | 2056 | 100.0 | 395 | 5 | ABB1554 | | Aab81554 Human cor | XX | | PN | WO200106015-A1. | |
| 3 | 2056 | 100.0 | 395 | 5 | AAB15438 | | Aae15438 Human dru | XX | | XX | | |
| 4 | 2056 | 100.0 | 395 | 7 | ADL21086 | | Adi21086 Novel hum | XX | | PD | 25-JAN-2001. | |
| 5 | 2056 | 100.0 | 395 | 8 | ADL61235 | | Adi61235 Human tyr | XX | | XX | 19-JUL-2000; 2000WO-US019741. | |
| 6 | 1934 | 94.1 | 395 | 5 | ABB81555 | | Aab81555 Consensus | XX | | PR | 20-JUL-1999; 99US-0144694P. | |
| 7 | 1729 | 84.1 | 390 | 4 | AYT2639 | | Aay72639 Human gly | XX | | PR | 13-JUN-2000; 2000US-00593828. | |
| 8 | 1729 | 84.1 | 390 | 5 | ABB81556 | | Aab81556 Human int | XX | | XX | (REGC) UNIV CALIFORNIA. | |
| 9 | 1708 | 83.1 | 395 | 4 | AYT2638 | | Aay72638 Mouse gly | PA | | | | |
| 10 | 1708 | 83.1 | 395 | 5 | AAU11275 | | Aau11275 Murine in | PA | | | | |
| 11 | 1694 | 82.4 | 418 | 5 | ABB81557 | | Aab81557 Mouse int | XX | | | | |
| 12 | 1620 | 78.8 | 418 | 3 | AAB41947 | | Aab41947 Human ORF | PI | | | | |
| 13 | 1538 | 74.8 | 394 | 7 | ADJ70405 | | Adj70405 Human hea | XX | | | | |
| 14 | 1024 | 49.8 | 386 | 4 | AAM93309 | | Aam93309 Human pol | XX | | | | |
| 15 | 1024 | 49.8 | 386 | 8 | ADL30784 | | Adl30784 Human pro | DR | | | | |
| 16 | 1019 | 49.6 | 386 | 2 | AYA39918 | | Aay39918 Human gly | XX | | | | |
| 17 | 1018 | 49.5 | 380 | 5 | AAU11274 | | Aau11274 Human L-s | PT | | | | |
| 18 | 981 | 47.7 | 386 | 3 | AYT2619 | | Aay72619 Human tra | XX | | | | |
| 19 | 977 | 47.5 | 388 | 2 | AYA39919 | | Aay39919 Mouse gly | XX | | | | |
| 20 | 808 | 39.3 | 169 | 5 | ABB81558 | | Aab81558 Human cor | PS | | | | |
| 21 | 755 | 36.7 | 169 | 5 | ABB81559 | | Aab81559 Human int | XX | | | | |
| 22 | 663 | 32.2 | 483 | 2 | AYA31656 | | Aay31656 Mouse N-a | CC | | | | |
| 23 | 657.5 | 32.0 | 530 | 4 | AAB93467 | | Aab93467 Human pro | CC | | | | |
| 24 | 657.5 | 32.0 | 530 | 8 | ADQ18590 | | Adq18590 Human sof | CC | | | | |
| 25 | 653.5 | 31.8 | 484 | 2 | AYA31657 | | Aay31657 Human N-a | CC | | | | |

ALIGNMENTS

RESULT 1
ID AAY72640 standard; protein: 395 AA.
XX AAY72640;
AC XX DT 02-MAY-2001 (first entry)

Human glycosyl sulfotransferase-4beta (GST-4beta).
DE Human glycosyl sulfotransferase-4beta (GST-4beta).
XX Human; glycosyl sulfotransferase-4beta; GST-4beta; immunosuppressive; inflammatory; therapy; selectin binding inhibitor; gene therapy; inflammation; systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes; polyarthritis nodosa; polymyositis; systemic sclerosis; dermatitis; glomerulonephritis; myasthenia gravis; Sjogren's syndrome; hypoparathyroidism; anaemia; Hashimoto's disease; Grise's disease; ulcerative colitis; allergic rhinitis; myocarditis; adult respiratory distress syndrome; eczema; psoriasis; asthma; hyperensitivity; rheumatic fever; tissue rejection; chromosome 16q23.1.

XX Homo sapiens.

OS XX PN WO200106015-A1.

XX PD 25-JAN-2001.

XX XX 19-JUL-2000; 2000WO-US019741.

XX XX 20-JUL-1999; 99US-0144694P.

XX PR 13-JUN-2000; 2000US-00593828.

XX XX (REGC) UNIV CALIFORNIA.

XX DR WPI; 2001-138471/14.

XX DR N-PSDB; AAD02697, AAD02700.

Rosen SD, Lee JK, Hemmerich S; New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications.

Claim 3, Fig 4B; 128pp; English.

The present sequence is human glycosyl sulfotransferase-beta (GST-4 beta). GST-4 Gene is found on chromosome 16q23.1. GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non

KW cancer; endocrine disorder; hypothalamus disorder; pituitary disorder;
 KW gastrointestinal disorder; metabolic disorder; developmental disorder;
 KW liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;
 KW DME-5.
 XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Peptide 1..35 /label= Signal_peptide
 FT Peptide 1..32 /label= Signal_peptide
 FT Protein 33..395 /note= "Human mature DME-5 protein"
 FT Protein 36..395 /note= "Human mature DME-5 protein"
 XX W0200179468-A2.
 XX PD 25-OCT-2001.
 XX PP 12-APR-2001; 2001WO-US011869.
 XX PR 13-APR-2000; 2000US-0197590P.
 PR 19-APR-2000; 2000US-0198403P.
 PR 28-APR-2000; 2000US-0200185P.
 PR 05-MAY-2000; 2000US-0202234P.
 PR 11-MAY-2000; 2000US-0203509P.
 XX PA (INCYT-) INCYTE GENOMICS INC.
 PI Policky JL, Hafalia A, Burford N, Ring HZ, Lai P, Tribouley CM;
 Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS;
 Gandhi AR, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA;
 PI Au-Young J;
 XX DR WPI; 2002-06363/09.
 DR N-PSDB; AAD24670.
 XX Novel isolated human drug metabolizing enzymes referred as DME 1-10,
 PT useful for diagnosing, treating, or preventing disorders associated with
 PT aberrant expression of DME such as allergy, anemia, asthma, infertility;
 XX PS Claim 1a: Page 131-132; 143PP; English.
 XX The invention relates to human drug metabolising enzymes referred as DME 1-10,
 CC and nucleic acid molecules encoding such enzymes. Polynucleotides of the
 CC invention are useful for assessing toxicity of test compounds and in gene
 CC therapy. Sequences of the invention are useful in the diagnosis,
 CC prevention and treatment of autoimmune/inflammatory disorders such as
 CC acquired immune deficiency syndrome (AIDS), adult respiratory distress
 CC syndrome, allergies, anaemia, atherosclerosis, asthma, autoimmune
 CC haemolytic anaemia, contact dermatitis, Crohn's disease,
 CC glomerulonephritis, Goodpasture's syndrome, gout, Grave's disease,
 CC Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,
 CC osteoarthritis, osteoporosis, psoriasis, systemic lupus erythematosus,
 CC rheumatoid arthritis, scleroderma, ulcerative colitis, uveitis, viral,
 CC bacterial, fungal, parasitic protozoal, helminthic infections, cell
 CC proliferative disorders such as actinic keratosis, arteriosclerosis,
 CC atherosclerosis, Duchenne and Becker dystrophy, hepatitis, Cushing's
 CC syndrome, cancers, myelodysplastic syndrome, epilepsy, endocrine
 CC disorders such as disorders of the hypothalamus and pituitary resulting
 CC from lesions such as primary brain tumours, adenomas, infarction
 CC associated with pregnancy, aneurysms, vascular malformations; eye
 CC disorders such as conjunctivitis, iritis, retinitis, glaucoma, pigmentosa
 CC ; metabolic disorders such as Addison's disease, cystic fibrosis,
 CC diabetes, goitre, glycogen storage diseases, hypercholesterolemia,
 CC mannosidosis, obesity; Gastrointestinal disorders such as dysphagia,
 CC gastric carcinoma, anoxia, nausea, gastroenteritis,
 CC hyperbilirubinaemia, emesis, cirrhosis, diarrhoea, jaundice, Reye's
 CC syndrome, peliosis hepatitis, hepatic vein thrombosis and developmental

CC disorders. The present sequence is human DME-5 protein
 XX SQ Sequence 395 AA;
 Query Match 100.0%; Score 2056; DB 5; Length 395;
 Best Local Similarity 100.0%; Pred. No. 3..1e-207;
 Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MMLPRVSTATVALLLQTFLPLFLYSPRGPSPAGEARVTVLVLSSWRSSSSFVGOLF 60
 Db 1 MMLPRVSTATVALLLQTFLPLFLYSPRGPSPAGEARVTVLVLSSWRSSSSFVGOLF 60
 Qy 61 NQHPDVYLMEPAWHVVTLTSQGSAATLHMAVRLDVLRSVFLCDMDVFLDAYLPMRNNSDL 120
 Db 61 NQHPDVYLMEPAWHVVTLTSQGSAATLHMAVRLDVLRSVFLCDMDVFLDAYLPMRNNSDL 120
 Qy 121 PQAVASRALCSPACSFPRGAISSEAVCKPLCARQSTFLAREACRSYSHVVLKEVRFFN 180
 Db 121 PQAVASRALCSPACSFPRGAISSEAVCKPLCARQSTFLAREACRSYSHVVLKEVRFFN 180
 Qy 181 LQVLYPLISDPAINLRLVHLVEDPRAVLRSRBOATAKALARDNGIVLGNGTWEDPLR 240
 Db 181 LQVLYPLISDPAINLRLVHLVEDPRAVLRSRBOATAKALARDNGIVLGNGTWEDPLR 240
 Qy 241 VYREVRSHVRLIAEATLKPPPLFRGRYRLVRFEDLAREPLAEIRALYAFTGSLSTPQLS 300
 Db 241 VYREVRSHVRLIAEATLKPPPLFRGRYRLVRFEDLAREPLAEIRALYAFTGSLSTPQLS 300
 Qy 301 AWIHNITHGSGPGEARAFKTSSRNALNVSQMRHALPFAKIRRVOELCAGAL-QLLGYRP 360
 Db 301 AWIHNITHGSGPGEARAFKTSSRNALNVSQMRHALPFAKIRRVOELCAGAL-QLLGYRP 360
 Qy 361 VYSEBDEQRNLALDVLVPLGLNGFTWAASSASHPRN 395
 Db 361 VYSEBDEQRNLALDVLVPLGLNGFTWAASSASHPRN 395
 RESULT 4
 AD121086
 ID AD121086 standard; protein: 395 AA.
 XX AC AD121086;
 XX DT 15-APR-2004 (first entry)
 XX DE Novel human protein #61.
 XX KW forensic; nutritional source; damaged tissue; diseased tissue;
 XX KW myeloid cell disorder; lymphoid cell disorder;
 XX KW bone cartilage tissue growth; tendon tissue growth;
 XX KW ligament tissue growth; nerve tissue growth; wound healing;
 XX KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
 OS Homo sapiens.
 XX WO2003025148-A2.
 XX PR 27-MAR-2003.
 XX PF 19-SEP-2002; 2002WO-US029964.
 XX PR 19-SEP-2001; 2001US-0323739P.
 PR 13-SEP-2002; 2002US-0323739.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
 PI Haley-Vicente D;
 DR WPI; 2003-354603/33.
 DR N-PSDB; ADI21802.

PT New polynucleotides and secreted proteins, useful for treating myeloid or
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
 PT tissue growth or regeneration, in wound healing, and in tissue repair and
 PT replacement.
 XX
 PS Claim 20; SEQ ID NO 337; 156pp; English.
 XX
 The invention relates to an isolated polynucleotide encoding a
 CC polypeptide with biological activity. The polynucleotides and
 CC polypeptides are useful in diagnostics, forensics, gene mapping,
 CC identification of mutations responsible for genetic disorders and other
 CC traits to assess biodiversity, as nutritional sources or supplements.
 CC The polynucleotides may also be used as molecular weight markers,
 CC chromosome markers or map related gene positions, or as an antigen to
 CC raise anti-DNA antibodies or elicit immune response. The polynucleotides are
 CC useful for raising antibodies, as markers for tissues in which the
 CC corresponding polypeptide is expressed, for re-engineering damaged or
 CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
 CC bone cartilage, tendon, ligament and/or nerve tissue growth or
 CC regeneration, in wound healing, in tissue repair and replacement, in
 CC healing of burns, incisions and ulcers, and in treating cancer. The
 CC present sequence represents the amino acid sequence of a novel human
 CC protein.
 XX
 SQ Sequence 395 AA;

Query Match 100.0%; Score 2056; DB 7; Length 395;
 Best Local Similarity 100.0%; Pred. No. 3.1e-207;
 Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLPRVSSATAVLLAQTFLFLFLVSPGPSSPAGGEARAVHVLSSWRSGSSFVQQLP 60
 Db 1 MNLPRVSSATAVLLAQTFLFLFLVSPGPSSPAGGEARAVHVLSSWRSGSSFVQQLP 60
 Qy 61 NQHDVFYLMEPAHVWWTLSQSAAATHMMAVDLVRSVFLCDMDVFDAYLPWRNISDL 120
 Db 61 NQHDVFYLMEPAHVWWTLSQSAAATHMMAVDLVRSVFLCDMDVFDAYLPWRNISDL 120
 Qy 121 FQWAVSRALCSPACSAFPGRASSEAVCKPLCARQSTTLAREACRSYSHVILKEYRFFN 180
 Db 121 FQWAVSRALCSPACSAFPGRASSEAVCKPLCARQSTTLAREACRSYSHVILKEYRFFN 180
 Qy 181 LQVLYPLISDPALNRIVHLVRDRAVLRSLRECTAKALARDNCIVLTGNTWEADGLR 240
 Db 181 LQVLYPLISDPALNRIVHLVRDRAVLRSLRECTAKALARDNCIVLTGNTWEADGLR 240
 Qy 241 VREVCRSHVRIAEEATLKPPFLRGYRLVRPDLAREPLAIRALYAFTGSLTQLE 300
 Db 241 VREVCRSHVRIAEEATLKPPFLRGYRLVRPDLAREPLAIRALYAFTGSLTQLE 300
 Qy 301 AWIHNITHSGGPGRARAEFKTSRNALNVSQVRHALPFAKIRRVOELCAGALQIGYRP 360
 Db 301 AWIHNITHSGGPGRARAEFKTSRNALNVSQVRHALPFAKIRRVOELCAGALQIGYRP 360
 Qy 361 VYSEDEQRMLALDVLPLCNGNTFWASSTASHRN 395
 Db 361 VYSEDEQRMLALDVLPLCNGNTFWASSTASHRN 395
 Qy 395
 Db 395
 Qy 400
 Db 400
 Qy 405
 Db 405
 Qy 410
 Db 410
 Qy 415
 Db 415
 Qy 420
 Db 420
 Qy 425
 Db 425
 Qy 430
 Db 430
 Qy 435
 Db 435
 Qy 440
 Db 440
 Qy 445
 Db 445
 Qy 450
 Db 450
 Qy 455
 Db 455
 Qy 460
 Db 460
 Qy 465
 Db 465
 Qy 470
 Db 470
 Qy 475
 Db 475
 Qy 480
 Db 480
 Qy 485
 Db 485
 Qy 490
 Db 490
 Qy 495
 Db 495
 Qy 500
 Db 500
 Qy 505
 Db 505
 Qy 510
 Db 510
 Qy 515
 Db 515
 Qy 520
 Db 520
 Qy 525
 Db 525
 Qy 530
 Db 530
 Qy 535
 Db 535
 Qy 540
 Db 540
 Qy 545
 Db 545
 Qy 550
 Db 550
 Qy 555
 Db 555
 Qy 560
 Db 560
 Qy 565
 Db 565
 Qy 570
 Db 570
 Qy 575
 Db 575
 Qy 580
 Db 580
 Qy 585
 Db 585
 Qy 590
 Db 590
 Qy 595
 Db 595
 Qy 600
 Db 600
 Qy 605
 Db 605
 Qy 610
 Db 610
 Qy 615
 Db 615
 Qy 620
 Db 620
 Qy 625
 Db 625
 Qy 630
 Db 630
 Qy 635
 Db 635
 Qy 640
 Db 640
 Qy 645
 Db 645
 Qy 650
 Db 650
 Qy 655
 Db 655
 Qy 660
 Db 660
 Qy 665
 Db 665
 Qy 670
 Db 670
 Qy 675
 Db 675
 Qy 680
 Db 680
 Qy 685
 Db 685
 Qy 690
 Db 690
 Qy 695
 Db 695
 Qy 700
 Db 700
 Qy 705
 Db 705
 Qy 710
 Db 710
 Qy 715
 Db 715
 Qy 720
 Db 720
 Qy 725
 Db 725
 Qy 730
 Db 730
 Qy 735
 Db 735
 Qy 740
 Db 740
 Qy 745
 Db 745
 Qy 750
 Db 750
 Qy 755
 Db 755
 Qy 760
 Db 760
 Qy 765
 Db 765
 Qy 770
 Db 770
 Qy 775
 Db 775
 Qy 780
 Db 780
 Qy 785
 Db 785
 Qy 790
 Db 790
 Qy 795
 Db 795
 Qy 800
 Db 800
 Qy 805
 Db 805
 Qy 810
 Db 810
 Qy 815
 Db 815
 Qy 820
 Db 820
 Qy 825
 Db 825
 Qy 830
 Db 830
 Qy 835
 Db 835
 Qy 840
 Db 840
 Qy 845
 Db 845
 Qy 850
 Db 850
 Qy 855
 Db 855
 Qy 860
 Db 860
 Qy 865
 Db 865
 Qy 870
 Db 870
 Qy 875
 Db 875
 Qy 880
 Db 880
 Qy 885
 Db 885
 Qy 890
 Db 890
 Qy 895
 Db 895
 Qy 900
 Db 900
 Qy 905
 Db 905
 Qy 910
 Db 910
 Qy 915
 Db 915
 Qy 920
 Db 920
 Qy 925
 Db 925
 Qy 930
 Db 930
 Qy 935
 Db 935
 Qy 940
 Db 940
 Qy 945
 Db 945
 Qy 950
 Db 950
 Qy 955
 Db 955
 Qy 960
 Db 960
 Qy 965
 Db 965
 Qy 970
 Db 970
 Qy 975
 Db 975
 Qy 980
 Db 980
 Qy 985
 Db 985
 Qy 990
 Db 990
 Qy 995
 Db 995
 Qy 1000
 Db 1000
 Qy 1005
 Db 1005
 Qy 1010
 Db 1010
 Qy 1015
 Db 1015
 Qy 1020
 Db 1020
 Qy 1025
 Db 1025
 Qy 1030
 Db 1030
 Qy 1035
 Db 1035
 Qy 1040
 Db 1040
 Qy 1045
 Db 1045
 Qy 1050
 Db 1050
 Qy 1055
 Db 1055
 Qy 1060
 Db 1060
 Qy 1065
 Db 1065
 Qy 1070
 Db 1070
 Qy 1075
 Db 1075
 Qy 1080
 Db 1080
 Qy 1085
 Db 1085
 Qy 1090
 Db 1090
 Qy 1095
 Db 1095
 Qy 1100
 Db 1100
 Qy 1105
 Db 1105
 Qy 1110
 Db 1110
 Qy 1115
 Db 1115
 Qy 1120
 Db 1120
 Qy 1125
 Db 1125
 Qy 1130
 Db 1130
 Qy 1135
 Db 1135
 Qy 1140
 Db 1140
 Qy 1145
 Db 1145
 Qy 1150
 Db 1150
 Qy 1155
 Db 1155
 Qy 1160
 Db 1160
 Qy 1165
 Db 1165
 Qy 1170
 Db 1170
 Qy 1175
 Db 1175
 Qy 1180
 Db 1180
 Qy 1185
 Db 1185
 Qy 1190
 Db 1190
 Qy 1195
 Db 1195
 Qy 1200
 Db 1200
 Qy 1205
 Db 1205
 Qy 1210
 Db 1210
 Qy 1215
 Db 1215
 Qy 1220
 Db 1220
 Qy 1225
 Db 1225
 Qy 1230
 Db 1230
 Qy 1235
 Db 1235
 Qy 1240
 Db 1240
 Qy 1245
 Db 1245
 Qy 1250
 Db 1250
 Qy 1255
 Db 1255
 Qy 1260
 Db 1260
 Qy 1265
 Db 1265
 Qy 1270
 Db 1270
 Qy 1275
 Db 1275
 Qy 1280
 Db 1280
 Qy 1285
 Db 1285
 Qy 1290
 Db 1290
 Qy 1295
 Db 1295
 Qy 1300
 Db 1300
 Qy 1305
 Db 1305
 Qy 1310
 Db 1310
 Qy 1315
 Db 1315
 Qy 1320
 Db 1320
 Qy 1325
 Db 1325
 Qy 1330
 Db 1330
 Qy 1335
 Db 1335
 Qy 1340
 Db 1340
 Qy 1345
 Db 1345
 Qy 1350
 Db 1350
 Qy 1355
 Db 1355
 Qy 1360
 Db 1360
 Qy 1365
 Db 1365
 Qy 1370
 Db 1370
 Qy 1375
 Db 1375
 Qy 1380
 Db 1380
 Qy 1385
 Db 1385
 Qy 1390
 Db 1390
 Qy 1395
 Db 1395
 Qy 1400
 Db 1400
 Qy 1405
 Db 1405
 Qy 1410
 Db 1410
 Qy 1415
 Db 1415
 Qy 1420
 Db 1420
 Qy 1425
 Db 1425
 Qy 1430
 Db 1430
 Qy 1435
 Db 1435
 Qy 1440
 Db 1440
 Qy 1445
 Db 1445
 Qy 1450
 Db 1450
 Qy 1455
 Db 1455
 Qy 1460
 Db 1460
 Qy 1465
 Db 1465
 Qy 1470
 Db 1470
 Qy 1475
 Db 1475
 Qy 1480
 Db 1480
 Qy 1485
 Db 1485
 Qy 1490
 Db 1490
 Qy 1495
 Db 1495
 Qy 1500
 Db 1500
 Qy 1505
 Db 1505
 Qy 1510
 Db 1510
 Qy 1515
 Db 1515
 Qy 1520
 Db 1520
 Qy 1525
 Db 1525
 Qy 1530
 Db 1530
 Qy 1535
 Db 1535
 Qy 1540
 Db 1540
 Qy 1545
 Db 1545
 Qy 1550
 Db 1550
 Qy 1555
 Db 1555
 Qy 1560
 Db 1560
 Qy 1565
 Db 1565
 Qy 1570
 Db 1570
 Qy 1575
 Db 1575
 Qy 1580
 Db 1580
 Qy 1585
 Db 1585
 Qy 1590
 Db 1590
 Qy 1595
 Db 1595
 Qy 1600
 Db 1600
 Qy 1605
 Db 1605
 Qy 1610
 Db 1610
 Qy 1615
 Db 1615
 Qy 1620
 Db 1620
 Qy 1625
 Db 1625
 Qy 1630
 Db 1630
 Qy 1635
 Db 1635
 Qy 1640
 Db 1640
 Qy 1645
 Db 1645
 Qy 1650
 Db 1650
 Qy 1655
 Db 1655
 Qy 1660
 Db 1660
 Qy 1665
 Db 1665
 Qy 1670
 Db 1670
 Qy 1675
 Db 1675
 Qy 1680
 Db 1680
 Qy 1685
 Db 1685
 Qy 1690
 Db 1690
 Qy 1695
 Db 1695
 Qy 1700
 Db 1700
 Qy 1705
 Db 1705
 Qy 1710
 Db 1710
 Qy 1715
 Db 1715
 Qy 1720
 Db 1720
 Qy 1725
 Db 1725
 Qy 1730
 Db 1730
 Qy 1735
 Db 1735
 Qy 1740
 Db 1740
 Qy 1745
 Db 1745
 Qy 1750
 Db 1750
 Qy 1755
 Db 1755
 Qy 1760
 Db 1760
 Qy 1765
 Db 1765
 Qy 1770
 Db 1770
 Qy 1775
 Db 1775
 Qy 1780
 Db 1780
 Qy 1785
 Db 1785
 Qy 1790
 Db 1790
 Qy 1795
 Db 1795
 Qy 1800
 Db 1800
 Qy 1805
 Db 1805
 Qy 1810
 Db 1810
 Qy 1815
 Db 1815
 Qy 1820
 Db 1820
 Qy 1825
 Db 1825
 Qy 1830
 Db 1830
 Qy 1835
 Db 1835
 Qy 1840
 Db 1840
 Qy 1845
 Db 1845
 Qy 1850
 Db 1850
 Qy 1855
 Db 1855
 Qy 1860
 Db 1860
 Qy 1865
 Db 1865
 Qy 1870
 Db 1870
 Qy 1875
 Db 1875
 Qy 1880
 Db 1880
 Qy 1885
 Db 1885
 Qy 1890
 Db 1890
 Qy 1895
 Db 1895
 Qy 1900
 Db 1900
 Qy 1905
 Db 1905
 Qy 1910
 Db 1910
 Qy 1915
 Db 1915
 Qy 1920
 Db 1920
 Qy 1925
 Db 1925
 Qy 1930
 Db 1930
 Qy 1935
 Db 1935
 Qy 1940
 Db 1940
 Qy 1945
 Db 1945
 Qy 1950
 Db 1950
 Qy 1955
 Db 1955
 Qy 1960
 Db 1960
 Qy 1965
 Db 1965
 Qy 1970
 Db 1970
 Qy 1975
 Db 1975
 Qy 1980
 Db 1980
 Qy 1985
 Db 1985
 Qy 1990
 Db 1990
 Qy 1995
 Db 1995
 Qy 2000
 Db 2000
 Qy 2005
 Db 2005
 Qy 2010
 Db 2010
 Qy 2015
 Db 2015
 Qy 2020
 Db 2020
 Qy 2025
 Db 2025
 Qy 2030
 Db 2030
 Qy 2035
 Db 2035
 Qy 2040
 Db 2040
 Qy 2045
 Db 2045
 Qy 2050
 Db 2050
 Qy 2055
 Db 2055
 Qy 2060
 Db 2060
 Qy 2065
 Db 2065
 Qy 2070
 Db 2070
 Qy 2075
 Db 2075
 Qy 2080
 Db 2080
 Qy 2085
 Db 2085
 Qy 2090
 Db 2090
 Qy 2095
 Db 2095
 Qy 2100
 Db 2100
 Qy 2105
 Db 2105
 Qy 2110
 Db 2110
 Qy 2115
 Db 2115
 Qy 2120
 Db 2120
 Qy 2125
 Db 2125
 Qy 2130
 Db 2130
 Qy 2135
 Db 2135
 Qy 2140
 Db 2140
 Qy 2145
 Db 2145
 Qy 2150
 Db 2150
 Qy 2155
 Db 2155
 Qy 2160
 Db 2160
 Qy 2165
 Db 2165
 Qy 2170
 Db 2170
 Qy 2175
 Db 2175
 Qy 2180
 Db 2180
 Qy 2185
 Db 2185
 Qy 2190
 Db 2190
 Qy 2195
 Db 2195
 Qy 2200
 Db 2200
 Qy 2205
 Db 2205
 Qy 2210
 Db 2210
 Qy 2215
 Db 2215
 Qy 2220
 Db 2220
 Qy 2225
 Db 2225
 Qy 2230
 Db 2230
 Qy 2235
 Db 2235
 Qy 2240
 Db 2240
 Qy 2245
 Db 2245
 Qy 2250
 Db 2250
 Qy 2255
 Db 2255
 Qy 2260
 Db 2260
 Qy 2265
 Db 2265
 Qy 2270
 Db 2270
 Qy 2275
 Db 2275
 Qy 2280
 Db 2280
 Qy 2285
 Db 2285
 Qy 2290
 Db 2290
 Qy 2295
 Db 2295
 Qy 2300
 Db 2300
 Qy 2305
 Db 2305
 Qy 2310
 Db 2310
 Qy 2315
 Db 2315
 Qy 2320
 Db 2320
 Qy 2325
 Db 2325
 Qy 2330
 Db 2330
 Qy 2335
 Db 2335
 Qy 2340
 Db 2340
 Qy 2345
 Db 2345
 Qy 2350
 Db 2350
 Qy 2355
 Db 2355
 Qy 2360
 Db 2360
 Qy 2365
 Db 2365
 Qy 2370
 Db 2370
 Qy 2375
 Db 2375
 Qy 2380
 Db 2380
 Qy 2385
 Db 2385
 Qy 2390
 Db 2390
 Qy 2395
 Db 2395
 Qy 2400
 Db 2400
 Qy 2405
 Db 2405
 Qy 2410
 Db 2410
 Qy 2415
 Db 2415
 Qy 2420
 Db 2420
 Qy 2425
 Db 2425
 Qy 2430
 Db 2430
 Qy 2435
 Db 2435
 Qy 2440
 Db 2440
 Qy 2445
 Db 2445
 Qy 2450
 Db 2450
 Qy 2455
 Db 2455
 Qy 2460
 Db 2460
 Qy 2465
 Db 2465
 Qy 2470
 Db 2470
 Qy 2475
 Db 2475
 Qy 2480
 Db 2480
 Qy 2485
 Db 2485
 Qy 2490
 Db 2490
 Qy 2495
 Db 2495
 Qy 2500
 Db 2500
 Qy 2505
 Db 2505
 Qy 2510
 Db 2510
 Qy 2515
 Db 2515
 Qy 2520
 Db 2520
 Qy 2525
 Db 2525
 Qy 2530
 Db 2530
 Qy 2535
 Db 2535
 Qy 2540
 Db 2540
 Qy 2545
 Db 2545
 Qy 2550
 Db 2550
 Qy 2555
 Db 2555
 Qy 2560
 Db 2560
 Qy 2565
 Db 2565
 Qy 2570
 Db 2570
 Qy 2575
 Db 2575
 Qy 2580
 Db 2580
 Qy 2585
 Db 2585
 Qy 2590
 Db 2590
 Qy 2595
 Db 2595
 Qy 2600
 Db 2600
 Qy 2605
 Db 2605
 Qy 2610
 Db 2610
 Qy 2615
 Db 2615
 Qy 2620
 Db 2620
 Qy 2625
 Db 2625
 Qy 2630
 Db 2630
 Qy 2635
 Db 2635
 Qy 2640
 Db 2640
 Qy 2645
 Db 2645
 Qy 2650
 Db 2650
 Qy 2655
 Db 2655
 Qy 2660
 Db 2660
 Qy 2665
 Db 2665
 Qy 2670
 Db 2670
 Qy 2675
 Db 2675
 Qy 2680
 Db 2680
 Qy 2685
 Db 2685
 Qy 2690
 Db 2690
 Qy 2695
 Db 2695
 Qy 2700
 Db 2700
 Qy 2705
 Db 2705
 Qy 2710
 Db 2710
 Qy 2715
 Db 2715
 Qy 2720
 Db 2720
 Qy 2725
 Db 2725
 Qy 2730
 Db 2730
 Qy 2735
 Db 2735
 Qy 2740
 Db 2740
 Qy 2745
 Db 2745
 Qy 2750
 Db 2750
 Qy 2755
 Db 2755
 Qy 2760
 Db 2760
 Qy 2765
 Db 2765
 Qy 2770
 Db 2770
 Qy 2775
 Db 2775
 Qy 2780
 Db 2780
 Qy 2785
 Db 2785
 Qy 2790
 Db 2790
 Qy 2795
 Db 2795
 Qy 2800
 Db 2800
 Qy 2805
 Db 2805
 Qy 2810
 Db 2810
 Qy 2815
 Db 2815
 Qy 2820
 Db 2820
 Qy 2825
 Db 2825
 Qy 2830
 Db 2830
 Qy 2835
 Db 2835
 Qy 2840
 Db 2840
 Qy 2845
 Db 2845
 Qy 2850
 Db 2850
 Qy 2855
 Db 2855
 Qy 2860
 Db 2860
 Qy 2865
 Db 2865
 Qy 2870
 Db 2870
 Qy 2875
 Db 2875
 Qy 2880
 Db 2880
 Qy 2885
 Db 2885
 Qy 2890
 Db 2890
 Qy 2895
 Db 2895
 Qy 2900
 Db 2900
 Qy 2905
 Db 2905
 Qy 2910
 Db 2910
 Qy 2915
 Db 2915
 Qy 2920
 Db 2920
 Qy 2925
 Db 2925
 Qy 2930
 Db 2930
 Qy 2935
 Db 2935
 Qy 2940
 Db 2940
 Qy 2945
 Db 2945
 Qy 2950
 Db 2950
 Qy 2955
 Db 2955
 Qy 2960
 Db 2960
 Qy 2965
 Db 2965
 Qy 2970
 Db 2970
 Qy 2975
 Db 2975
 Qy 2980
 Db 2980
 Qy 2985
 Db 2985
 Qy 2990
 Db 2990
 Qy 2995
 Db 2995
 Qy 3000
 Db 3000
 Qy 3005
 Db 3005
 Qy 3010
 Db 3010
 Qy 3015
 Db 3015
 Qy 3020
 Db 3020
 Qy 3025
 Db 3025
 Qy 3030
 Db 3030
 Qy 3035
 Db 3035
 Qy 3040
 Db 3040
 Qy 3045
 Db 3045
 Qy 3050
 Db 3050
 Qy 3055
 Db 3055
 Qy 3060
 Db 3060
 Qy 3065
 Db 3065
 Qy 3070
 Db 3070
 Qy 3075
 Db 3075
 Qy 3080
 Db 3080
 Qy 3085
 Db 3085
 Qy 3090
 Db 3090
 Qy 3095
 Db 3095
 Qy 3100
 Db 3100
 Qy 3105
 Db 3105
 Qy 3110
 Db 3110
 Qy 3115
 Db 3115
 Qy 3120
 Db 3120
 Qy 31

ABB81555
 ID ABB81555 standard; protein; 395 AA.
 XX
 AC ABB81555;
 XX
 DT 05-SEP-2002 (first entry)
 XX
Consensus N-acetylglucosamine-6-sulfotransferase SEQ ID NO:3.
 DE Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GICNA6ST;
 KW corneal; sulfatation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.
 XX
 OS Homo sapiens.
 OS Mus musculus.
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 10 /label= Ala, Thr, Val
 FT Misc-difference 13 /label= Ala, Val, Ser
 FT Misc-difference 14 /label= Ala, Val, Ser
 FT Misc-difference 20 /label= Phe, Cys, Gly
 FT Misc-difference 39 /label= Ala, Asp, Glu
 FT Misc-difference 96 /label= Val, Met, Ile
 FT Misc-difference 142 /label= Ala, Thr, Asn
 FT Misc-difference 147 /label= Ala, Asp, Glu
 FT Misc-difference 159 /label= Ala, Asp, Glu
 FT Misc-difference 238 /label= Thr, Ser, Gly
 FT Misc-difference 294 /label= Gly, His, Arg
 FT Misc-difference 371 /label= Ser, Thr, Gly
 FT Misc-difference 380 /label= Ala, Thr, Ser
 FT Misc-difference 382 /label= Leu, Pro, Met
 FT Misc-difference 384 /label= Gly, His, Ser
 FT Misc-difference 390 /label= Thr, Ser, Lys
 FT Misc-difference 391 /label= Ala, Glu
 FT Misc-difference 395 /label= Asn, Ser, Lys
 FT Misc-difference 392 /label= His, Gln
 FT Misc-difference 394 /label= Arg, Glu
 FT Misc-difference 395 /label= Asn, Ser
 XX
 PN US2002061562-A1.
 XX
 PD 23-MAY-2002.
 XX
 PP 09-AUG-2001; 2001US-00927602.
 XX
 PR 11-AUG-2000; 2000US-00638211.
 PR 11-AUG-2000; 2000US-0325773P.
 PA (FUKU/) FUKUDA, M. N.
 PA (AKAM/) AKAMA T O.
 PI Fukuda MN, Akama TO;
 XX
 DR 2002-507643/54.

PT New nucleic acid encoding corneal N-acetylglucosamine-6-sulfotransferase,
 PT useful for treatment, monitoring and diagnosis of macular corneal
 PR dystrophy.
 XX
 PS Example 5; Fig 2A-B; 69pp; English.
 XX
 CC The present invention describes human corneal N-acetylglucosamine-6-
 CC sulfotransferase (I), which is able to catalyze sulfation of keratan
 CC sulfate (KS). Also, described is a method for monitoring the effect of
 CC treatments for macular corneal dystrophy (MCD), and detecting
 CC susceptibility to MCD. (I) is located to chromosome 16q22 and has
 CC ophthalmological activity. (I) can be used to treat or prevent macular
 CC corneal dystrophy types I or II. (I) makes possible treatment of MCD
 CC without requiring keratoplasty or keratectomy. The present sequence
 CC represents a consensus N-acetylglucosamine-6-sulfotransferase which is
 CC given in the exemplification of the present invention
 XX
 Sequence 395 AA:
 SQ Query Match 94.1%; Score 1934; DB 5; Length 395;
 Best Local Similarity 94.4%; Pred. No. 2.e-104;
 Matches 371; Conservative 3; Mismatches 19; Indels 0; Gaps 0;
 Qy 1 MNLPRVSSATAVALLAQTFLLFLVSRGPSSSPAGGBEARVNTLSSWRSGSFVGQFLP 60
 Db 1 MNLPRFSSSTXTVXLLAQTXLLFLVSRGPSSSPAGGEKRVHVLVLSWRSGSFVGQFLP 60
 Qy 61 NQPDPVYLMEPAWHVVTTLSQGSAAATHMAVRDLVSVFLCDMDVFLDAYLPRRNLSDL 120
 Db 61 SQPDPVYLMEPAWHVVTTLSQGSAAATHMAVRDLXSVFLCDMDVFLDAYLPRRNLSDL 120
 Qy 121 FQAVASRALCSPPACSAFPRGAISSAEVCKPLCARQSTTLAREACRSTSHTVKEVRPN 180
 Db 121 FQAVASRALCSPPACSAFPRGXISSEVCKPLCARQPTXLAREACRSTSRETAKALDNGIVLGTGNTWAEADPLIR 180
 Qy 181 LQVLYPLISDPALNLRVHLVRDPRAVRLSREDTAKALDNGIVLGTGNTWAEADPLIR 240
 Db 181 LQVLYPLISDPALNLRVHLVRDPRAVRLSREDTAKALDNGIVLGTGNTWAEADPLIR 240
 Qy 241 VREVCRHVRIAEATLKPPPLRGRYRLVRFEDLAREPLAIIRALYAFTGSLSTLPQIE 300
 Db 241 VREVCRHVRIAEATLKPPPLRGRYRLVRFEDLAREPLAIIRALYAFTGSLXLTLPQIE 300
 Qy 301 AWIHNITIGSGPARRAKFTTSRNALNVSQWRLPAFAKTRVQLCAGALQLLGYRP 360
 Db 301 AWIHNITIGSGPARRAKFTTSRNALNVSQWRLPAFAKTRVQLCAGALQLLGYRP 360
 Qy 361 VYSEDQRLNDLVLPGLNGTWASSTASHP 393
 Db 361 VYSEDQRLNDLVLPGDXFEXWASSTXXXP 393
 AC AAV72639;
 XX
 RESULT 7
 AAV72639 standard; protein; 390 AA.
 XX
 DE Human glycosyl sulfotransferase-4alpha (GST-4alpha).
 XX
 KW therapy; selectin binding inhibitor; gene therapy; inflammation;
 KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
 KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
 KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenitis;
 KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
 KW denyeinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
 KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
 KW asthma; hypersensitivity; rheumatic fever; tissue rejection;
 KW chromosome 16q23.1.
 XX

| | | | |
|----------|---|---|-----|
| db | 181 | NLQVLYPLISDPALNRLTHVLDRPRAVRSREAGPIIARDGIVLGNGKWDADPHL | 240 |
| DY | 240 | RVREVCNSHVRIAEATKPPPLRGYRVLURPEDLAREPLAETRALYARTGLSLTPOL | 299 |
| db | 241 | RJJREVCNSHVRAEAATKPPPLRGYRVLURPEDLAREPLAETRALYARTGLSLTPOL | 300 |
| DY | 300 | EAWINITHGSIGPSARAEFKTSRNALNVSQMRHALPFAKIRRVQELCAGALQLLGVR | 359 |
| db | 301 | EAWINITHGSIGCKPIEAFTHSRNARNVSQMRHALPFKILRVQECAQALQLGTR | 360 |
| DY | 360 | PVYSEDEQRNLALDVLPRGLNGTWA | 367 |
| db | 361 | PVYSDAQDQRDLTDLVLPRGDHFWSWAS | 368 |
| RESULT 1 | | | |
| ID | AAY72638 | standard; protein; 395 AA. | |
| XX | AAY72638; | | |
| XX | 02-MAY-2001 | (first entry) | |
| XX | Mouse glycosyl sulfotransferase-4 (GST-4). | | |
| XX | Mouse; glycosyl sulfotransferase-4; GST-4; immunosuppressive; therapy; | | |
| XX | selectin binding inhibitor; gene therapy; inflammation; | | |
| XX | systemic lupus erythematosus (SLE); rheumatoid arthritis; diabetes; | | |
| XX | polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis; | | |
| XX | glomerulonephritis; myasthenia Gravis; Sjogren's syndrome; adrenalitis; | | |
| XX | Bashimoto's disease; Grave's disease; hypoparathyroidism; anaemia; | | |
| XX | demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis; | | |
| XX | myocarditis; adult respiratory distress syndrome; eczema; psoriasis; | | |
| XX | asthma; hypersensitivity; rheumatic fever; tissue rejection; | | |
| XX | chromosome 9E1. | | |
| XX | Mus musculus. | | |
| XX | W0200106015-A1. | | |
| XX | 25-JAN-2001. | | |
| XX | 19-JUL-2000; 2000WO-US019741. | | |
| PP | 20-JUL-1999; 99US-01446949. | | |
| PR | 13-JUN-2000; 2000US-00593828. | | |
| PR | (REGC) UNIV CALIFORNIA. | | |
| XX | Rosen SD, Lee JK, Hammerich S; | | |
| XX | WPI; 2001-138471/14. | | |
| DR | NP-PSDB; AAD02696. | | |
| XX | New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications. | | |
| XX | Claim 3; Fig 2; 128pp; English. | | |
| XX | The present sequence is mouse glycosyl sulfotransferase-4 (GST-4). GST-4 gene is found on chromosome 9E1. GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non-sulphated selectin ligand, GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in inhibiting a selectin mediated binding event. GST is useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis, Hashimoto's syndrome, Sjogren's syndrome, Hashimoto's disease, Grave's disease, demyelinating diseases, Cirrhosis, ulcerative colitis, dermatitis, myasthenia gravis, hypothyroidism, pernicious anemia, Graves' disease, renal artery stenosis, psoriasis, lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation | | |
| XX | Sequence 395 AA: | | |
| XX | Query Match 83.1%; Score 1708; DB 4; Length 395; | | |
| XX | Best Local Similarity 83.0%; Pred. No. 1-4e-170; | | |
| XX | Matches 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0; | | |
| XX | SQ | | |
| Qy | 1 MWLPRVSSATAVALLAQTTEFLFLVSRGPSPAGEARVHYLVLSMRSGSSFVGOLF 60 | | |
| XX | 1 MRLPRFESTVMSLMLMWTGILFLVLSQVSPSSAGLGRVHLVLSMRSGSSFVGOLF 60 | | |
| Db | 1 SQHPDVYLMEDAWHNTLTLQCSAATLHMAYVLDLVRVSYFLCDMDVFDAYLPWRRNLSDL 120 | | |
| Qy | 61 NOHPDVYLMEDAWHNTLTLQCSAATLHMAYVLDLVRVSYFLCDMDVFDAYLPWRRNLSDL 120 | | |
| Db | 61 SQHPDVYLMEDAWHNTLTLQCSAATLHMAYVLDLVRVSYFLCDMDVFDAYLPWRRNLSDL 120 | | |
| Qy | 62 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 62 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 63 FQAVSRLACSPiACSFPRGAISSEAVCKPLCARQSFTLAREACSYSHVVLKEVRFFN 180 | | |
| Db | 63 FQAVSRLACSPiACSFPRGAISSEAVCKPLCARQSFTLAREACSYSHVVLKEVRFFN 180 | | |
| Qy | 64 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 64 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 65 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 65 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 66 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 66 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 67 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 67 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 68 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 68 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 69 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 69 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 70 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 70 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 71 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 71 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 72 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 72 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 73 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 73 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 74 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 74 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 75 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 75 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 76 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 76 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 77 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 77 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 78 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 78 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 79 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 79 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 80 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 80 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 81 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 81 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 82 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 82 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 83 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 83 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 84 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 84 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 85 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 85 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 86 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 86 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 87 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 87 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 88 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 88 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 89 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 89 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 90 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 90 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 91 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 91 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 92 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 92 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 93 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 93 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 94 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 94 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 95 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 95 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 96 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 96 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 97 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 97 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 98 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 98 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 99 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 99 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 100 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 100 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 101 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 101 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 102 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 102 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 103 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 103 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 104 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 104 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 105 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 105 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 106 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 106 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 107 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 107 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 108 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 108 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 109 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 109 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 110 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 110 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 111 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 111 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 112 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 112 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 113 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 113 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 114 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 114 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 115 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 115 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 116 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 116 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 117 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 117 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 118 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 118 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 119 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 119 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 120 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 120 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 121 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 121 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 122 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 122 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 123 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 123 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 124 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 124 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 125 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 125 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 126 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 126 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 127 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 127 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 128 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 128 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 129 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 129 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 130 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 130 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 131 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 131 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 132 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 132 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 133 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 133 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 134 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 134 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 135 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 135 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 136 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 136 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 137 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 137 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 138 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Db | 138 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARDNGTVLGTNTWEDPGLR 240 | | |
| Qy | 139 1QDLYPLISDPAINLRVHLYDPRVATYRSRECATKALARD | | |

XX WPI; 2003-845369/78.
 DR PD 05-SEP-2001.
 XX PP 07-JUN-2000; 2000EP-00114089.
 PT PT 08-JUL-1999; 99JP-00194786.
 PT PR 11-JAN-2000; 2000JP-0011874.
 PT PR 02-MAY-2000; 2000JP-00183765.
 XX PA (HELI-) HELIX RES INST.
 XX Ota T, Nishikawa T, Isocai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 DR N-PSDB; AA#94229.
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PR in genetic manipulation.
 XX Claim 8; SEQ ID NO 2817; 1380pp + Sequence Listing; English.
 XX The invention relates to primers for synthesising full length cDNA.
 CC 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesising the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a polypeptide encoded by a full length
 CC human cDNA of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in CD-ROM
 XX format directly from EPO.

Sequence 394 AA;

| Query Match | Score 1538; | DB 7; | Length 394; | |
|---|--------------|------------|-------------|--|
| Best Local Similarity | 93.3%; | Pred. No. | 1.1e-152; | |
| Matches 294; | Conservative | Mismatches | 14; | |
| | Indels | Gaps | 0; | |
| Qy 1 MWLPRVSSTAVTALLAOTFLFLVSRPGPSSPAGGEARHVLVLSWRSGSSFYVGOLF 60 | | | | |
| Db 1 MWLPRVSSTAVTALLAOTFLFLVSRPGPSSPAGGEARHVLVLSWRSGSSFYVGOLF 60 | | | | |
| Qy 61 NQHPDVYLMPEAWHWTLSQSQAATLHMAYRDLYSVELCDMDVFDAYLPWRRNISDL 120 | | | | |
| Db 61 NQHPDVYLMPEAWHWTLSQSQAATLHMAYRDLYSVELCDMDVFDAYLPWRRNISDL 120 | | | | |
| Qy 121 FQWAVSRLCSPACSAFRGAISSAEAVCKPLCARQSTTLAREACRSYSHVILKEYFFN 180 | | | | |
| Db 121 FQWAVSRLCSPACSAFRGAISSAEAVCKPLCARQSTTLAREACRSYSHVILKEYFFN 180 | | | | |
| Qy 181 LQVLYPLSDPAINLRVTHLVRDRAVRSREAGAPILARDIGIVLTGNGKVEADGLR 240 | | | | |
| Db 181 LQVLYPLSDPAINLRVTHLVRDRAVRSREAGAPILARDIGIVLTGNGKVEADGLR 240 | | | | |
| Qy 241 VREVCRSHVIRIAEATLKP PPPFLGRYRLVRFEDLAREPLAEIRALYAFTGSLITPQLE 300 | | | | |
| Db 241 VREVCRSHVIRIAEATLKP PPPFLGRYRLVRFEDLAREPLAEIRALYAFTGSLITPQLE 300 | | | | |
| Qy 301 AWIHNITHGCGPBAR 315 | | | | |
| Db 301 AWIHNITHGCGPBAR 315 | | | | |

RESULT 14
 AAM93309 ID AAM93309 standard; protein; 386 AA.
 XX AC AAM93309;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human Polypeptide, SEQ ID NO: 2817.
 XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX OS Homo sapiens.
 XX PN EP1130094-A2.
 XX PN ADL30784 standard; protein; 386 AA.

Query Match 74.8%; Score 1538; DB 7; Length 394;
 Best Local Similarity 93.3%; Pred. No. 1.1e-152;
 Matches 294; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
 Sequence 386 AA;

| Query Match | Score 1024; | DB 4; | Length 386; |
|--|--------------|--------------------|-------------|
| Best Local Similarity | 54.8%; | Pred. No. 1.6e-38; | |
| Matches 217; | Conservative | Mismatches | 44; |
| | Indels | Gaps | 7; |
| Oy 14 LLLAQTFILLFLVSRP-----GPSSPAGGEARHVLVLSWRSGSSFYVGOLF 59 | | | |
| Db 2 LLPKKNKLFLVSRPQMAIALAFFMHYSHTNSSLSMKAQPERMHVLVLSWRSGSSFYVGOLF 61 | | | |
| Qy 60 FNQHPDVYLMPEAWHWTLSQSQAATLHMAYRDLYSVELCDMDVFDAYLPWRRNISDL - PWRNLN 118 | | | |
| Db 62 FGQHPDVYLMPEAWHWTLSQSQAATLHMAYRDLYSVELCDMDVFDAYLPWRRNISDL - PWRNLN 121 | | | |
| Qy 119 DLFQWAVSRLCSPACSAFRGAISSAEAVCKPLCARQSTTLAREPLAEIRALYAFTGSLITPQLE 178 | | | |
| Db 122 SLFWQENWSRLCSPACDIQDETPRAHCRLLCSQQPFEEVAKRSYSHVILKEYFFN 181 | | | |
| Qy 179 FNQVLYPLSDPAINLRVTHLVRDRAVRSREAGAPILARDIGIVLTGNGKVEADGLR 237 | | | |
| Db 182 FNQVLYPLSDPAINLRVTHLVRDRAVRSREAGAPILARDIGIVLTGNGKVEADGLR 241 | | | |
| Qy 238 GLRVYREVCRSHVIRIAEATLKP - PPFLGRYRLVRFEDLAREPLAEIRALYAFTGSLIT 296 | | | |
| Db 242 PYVMQVICQSQLEIYK--TQSPLPKALQERYLLYEDLARAPYAQTSRMVEFVGLF 299 | | | |
| Qy 297 PQEAEATHNITHGCGPBARFAEKTSRNALNVSQAWRHALPFAKTRVOPLCAGAQQLL 356 | | | |
| Db 300 PHLQTWVNITRGKGNG--DEAFTNARDALNVSQWRWSPYEKVRSLQACGDANLNL 357 | | | |
| Qy 357 GYRVYSEDEORNLAIDLVLRGLNGFTWASSTASH 392 | | | |
| Db 358 GYRHVSSEQEQRNLIDLL-----STWVPEQIH 386 | | | |

| | | | | | | | | |
|----|-------------------------------|--|-----|--|--|--|--|--|
| AC | ADL30784; | | | | | | | |
| XX | 20-MAY-2004 | (first entry) | | | | | | |
| DT | XX | Human protein encoded by a full length cDNA clone SegID 2817. | | | | | | |
| DE | XX | human; medicine; signal transduction; glycoprotein; transcription; oligo-capping method. | | | | | | |
| KW | XX | | | | | | | |
| KW | XX | | | | | | | |
| XX | OS | Homo sapiens. | | | | | | |
| XX | XX | | | | | | | |
| PN | XX | EPI396543-A2. | | | | | | |
| PD | XX | 10-MAR-2004. | | | | | | |
| XX | PP | 07-JUL-2000; 2003EP-00025638. | | | | | | |
| XX | PR | 08-JUL-1999; 99JP-00194486. | | | | | | |
| PR | 11-JAN-2000; 2000JP-00118774. | | | | | | | |
| PR | 02-MAY-2000; 2000JP-00183865. | | | | | | | |
| PR | 07-JUL-2000; 2000EP-0014089. | | | | | | | |
| XX | (REAS-) | RES ASSOC BIOTECHNOLOGY. | | | | | | |
| XX | PA | | | | | | | |
| PI | PI | Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y; PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H; | | | | | | |
| XX | DR | WPI: 2004-204755/20. | | | | | | |
| XX | DR | N-PSDB; ADL30783. | | | | | | |
| XX | PT | New oligonucleotide primers (830 cDNAs) useful for synthesizing full length human cDNAs. | | | | | | |
| XX | PT | Example 1; SEQ ID NO 2817; 1340pp; English. | | | | | | |
| XX | CC | This invention relates to a novel primers useful for synthesising full length cDNA molecules that encode human proteins. Specifically, it refers to secretory or membrane proteins that are potential therapeutic agents/ target molecules in the field of medicine, and in particular genes/ encoding proteins that are associated with signal transduction, glycoproteins and transcription. The present invention describes a method for efficiently cloning a full length human cDNA from both the 5' and 3' ends using the oligo-capping method. This polypeptide sequence is a full length human protein of the invention. | | | | | | |
| XX | SQ | Sequence 386 AA; | | | | | | |
| Qy | 14 | LLAQTEFLIFLFLYSPR-----GPSSPAGGEARVHLVLSWRSSESSSFVQL | 59 | | | | | |
| Db | 2 | LLPKRMKLUFLFYSQMATALLFFHMYSHNIISSLSMKAQFRMHVLVLSWRSSESSSFVQL | 61 | | | | | |
| Qy | 60 | FNQHPDPVFLMPEPAHWVWTLSQCSAATHMAYARDLVRSPFLCDMDVSDAYL_PWRNL | 118 | | | | | |
| Db | 62 | FGQHPDPVFLMPEPAHWVWTLSQCSAATHMAYARDLIRAVFLCDMSVTDAYMGPGRQS | 121 | | | | | |
| Qy | 119 | DLFQMAVSRLCSPPPACSAFPQRGAISSAECKPLCAROSFTLARBCHSYSHVLEKRF | 178 | | | | | |
| Db | 122 | SLFQENSRALCSAPACDITPQDELIIPRANCHLCSQQPFVEKACRSYSHVLEKRF | 181 | | | | | |
| Qy | 179 | FNQVQLYPLISDPAINLRYVHLVDPRAVRSRPTOKALARONGIVTGTW-EADP | 237 | | | | | |
| Db | 182 | FNQVQLYPLISDPAINLRYVHLVDPRAVRSRPTKGDLMIDRIVNGQHEQLKEDQ | 241 | | | | | |
| Qy | 238 | GLRVYREVCHSHVRIAEEAATLKP_PPFILAGRYLVRFPEDLAERPLAETRALYAFTLSLT | 296 | | | | | |
| Db | 242 | PYYMQVICCSQLEYK-TIQSULPKALQERYLIVREBLLARPAVQTSRMSFVGLEFL | 299 | | | | | |
| Qy | 297 | POLEARWHNTTHGSGPGARREAFKTSSRNALNYSQAWRHALPPAKIRREVQELCAGALOIL | 356 | | | | | |

THIS PAGE BLANK (USPTO)

| Result No. | Score | Query | Match | Length | DB ID | Description |
|------------|--------|-------|-------|--------|---------------------|-------------------|
| 1 | 2056 | 100.0 | 395 | 4 | US-09-949-016-7011 | Sequence 7011, Ap |
| 2 | 1729.5 | 84.1 | 390 | 4 | US-09-949-016-8813 | Sequence 6813, Ap |
| 3 | 1729.5 | 84.1 | 431 | 4 | US-09-949-016-8893 | Sequence 8893, Ap |
| 4 | 1019 | 49.6 | 386 | 3 | US-09-045-284A-2 | Sequence 2, Appl |
| 5 | 1019 | 49.6 | 386 | 3 | US-09-190-911-1 | Sequence 1, Appl |
| 6 | 981 | 47.7 | 386 | 4 | US-09-786-240-11 | Sequence 11, Appl |
| 7 | 663 | 32.2 | 483 | 3 | US-09-263-023-2 | Sequence 2, Appl |
| 8 | 663 | 32.2 | 483 | 4 | US-09-471-867-2 | Sequence 2, Appl |
| 9 | 657.5 | 32.0 | 608 | 4 | US-09-949-016-9449 | Sequence 9449, Ap |
| 10 | 653.5 | 31.8 | 458 | 2 | US-08-655-878-2 | Sequence 4, Appl |
| 11 | 653.5 | 31.8 | 484 | 4 | US-09-471-867-4 | Sequence 4, Appl |
| 12 | 653.5 | 31.8 | 531 | 4 | US-09-949-016-9471 | Sequence 6471, Ap |
| 13 | 576.5 | 28.0 | 479 | 2 | US-08-899-514-2 | Sequence 2, Appl |
| 14 | 555.5 | 27.0 | 411 | 3 | US-09-015-88-2 | Sequence 2, Appl |
| 15 | 532.5 | 25.9 | 458 | 2 | US-08-655-878-2 | Sequence 2, Appl |
| 16 | 100.5 | 4.9 | 199 | 4 | US-09-52-91A-28497 | Sequence 28497, A |
| 17 | 96 | 4.7 | 566 | 4 | US-08-484-93B-41 | Sequence 41, Appl |
| 18 | 96 | 4.7 | 566 | 2 | US-08-484-158B-41 | Sequence 41, Appl |
| 19 | 96 | 4.7 | 566 | 2 | US-08-484-596A-41 | Sequence 41, Appl |
| 20 | 96 | 4.7 | 566 | 2 | US-08-480-150A-41 | Sequence 41, Appl |
| 21 | 96 | 4.7 | 566 | 3 | US-08-458-731-41 | Sequence 41, Appl |
| 22 | 96 | 4.7 | 566 | 3 | US-08-149-723A-41 | Sequence 41, Appl |
| 23 | 95.5 | 4.6 | 268 | 4 | US-09-818-180-17 | Sequence 17, Appl |
| 24 | 95.5 | 4.6 | 268 | 4 | US-09-818-780-94 | Sequence 94, Appl |
| 25 | 93.5 | 4.5 | 575 | 4 | US-09-252-91A-32986 | Sequence 32986, A |
| 26 | 93.5 | 4.5 | 1385 | 2 | US-08-687-399-7 | Sequence 7, Appl |
| 27 | 93 | 4.5 | 359 | 3 | US-09-150-133-11 | Sequence 11, Appl |

RESULT 2

US-09-949-016-6813

; Sequence 6813, Application US/0949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 6813

; LENGTH: 390

; TYPE: PRT

; ORGANISM: Human

; US-09-949-016-6813

Query Match Score 84.1%; DB 4; Length 431;

Best Local Similarity 85.8%; Pred. No. 4.5e-175;

Matches 333; Conservative 17; Mismatches 37; Indels 1; Gaps 1;

Qy 1 MWLPRVSSTATAILAQ-TFLILFLYSPRGPSPAGEARVHVLSSWRSGSSFVGQL 59

Db 42 MWLPRFSKTVVLLAQTTCLLFIISRGPSPAGEARVHVLSSWRSGSSFVGQL 1.01

Qy 60 FNOHPDVFYLMEPAWHTWTLSQGSAATLHMAVRDPLRVSFLCDMDVDAVLPWRNLSD 119

Db 1.02 FSQHPDVFYLMEPAWHTWTLSQGSAATLHMAVRDPLRVSFLCDMDVDAVLPWRNLSD 161

Qy 120 LFQWAVSRLCSPACSAFPGRASSEAVCPLCAROSFTLAREACSYSHVVLKEVRFF 179

Db 1.62 FFWATRALSPPASSAFPCTISQDVCKTLCQPSLAREACSYSHVVLKEVRFF 221

Qy 180 NLQVLYPLISDPAINLRVHLRDPRAVLSREQTAKALARONGIVLTGNTGWVEADPGL 239

Db 222 NLQVLYPLISDPAINLRVHLRDPRAVLSREQTAKALARONGIVLTGNTGWVEADPGL 281

Qy 240 RVYREVRSHYRIAEATLKPPEFLGRGRYRLVRFEDAREPLAEIRALYAFTGIL-SITPQL 299

Db 282 RLIREVRSHYRIAEATLKPPEFLGRGRYRLVRFEDAREPLAEIRALYAFTGILTPQL 341

Qy 300 EAWHNTHGSGPGRARFAKTSRNMLNTSQAWRHALPFAKTRVQECLAGALQIGYR 359

Db 342 EAWHNTHGSGIIGKPEAFTSSRNMLNTSQAWRHALPFKTRVQECLAGALQIGYR 401

RESULT 4

US-09-045-284A-2

; Sequence 2, Application US/09045284A

; Patent No. 6255192

; GENERAL INFORMATION:

; APPLICANT: Bistrup, Annette

; APPLICANT: Rosen, Steven D.

; APPLICANT: Hemmerich, Stefan

; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3

; FILE REFERENCE: 6510-107US1

; CURRENT APPLICATION NUMBER: US/09/045-284A

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 2

; LENGTH: 386

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-045-284A-2

Query Match Score 1019; DB 3; Pred. No. 1.4e-99;

Best Local Similarity 54.5%;

US-09-045-284A-2

Query Match Score 1019; DB 3; Pred. No. 1.4e-99;

Best Local Similarity 54.5%;

RESULT 3

US-09-949-016-8893

; Sequence 8893, Application US/0949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

Matches 216; Conservative 45; Mismatches 107; Indels 28; Gaps 7;

Qy 14 LLLAQTFULLFLVSRP-----GPSSPAGGEARHVYLVLSSWRSGSFVGQL 59
Db 2 LLPKMKLFLFLVSMALLFFMYSNHSNISLMSMKAQPERMHLVLSWRSGSFVGQL 61

Qy 60 FNQHPDVYLMEPAWHWTLSQSGSAATLHMAYRDLVRSLVFLCDMDVDAYL_PWRNLIS 118
Db 62 FGQHPDVYLMEPAWHWTLSQSGSAATLHMAYRDLVRSLVFLCDMSVFDAYMGPPIQS 121

Qy 119 DLFOAVASRALCSPACSAFPRGAISSCAVKPLICAROSTTLEACRSYSHVILKEYRF 178
Db 122 SLFQWENSRALCSAPACDIPQDEIIPRAHCRLJCSCQPFVEKACSYSHVILKEYRP 181

Qy 179 FNQLYPLISDPALNLRIVHLVRDPRAVRSRETOAKALARDNGIVLTGNTW-EADP 237
Db 182 FNLSQYPLIKDPSLNHLTHVLRDPRAVRSRETKGDMLIDSRYINGQHEQKLKKEDQ 241

Qy 238 GLRVYREVCRSHVRIAEAATLKP_PPELRGRYLVRFEDLAREPLAETRALLYFTGLSLT 296
Db 242 PYVMQVICQSQLEIYK-TIQSLPKALQERYLVRLVRDALARPAVQTSRMYBEGGLEFL 299

Qy 297 POLEWIHNITHGSGP GARREAFKTSSENALNSQAMRHFAKIRRVQELCAGALQQL 356
Db 300 PHQOTWVNITRGKGMG-DHAFTNARDALNSQAMRWSLPEKVSRLQACGDMNL 357

Qy 357 GYRPVYSEDQRLALDVLVLRGNGFTWASSSTASH 392
Db 358 GYRHVRSEQRNLNDLL-----STWTVPEQIH 386

RESULT 6
US-09-786-240-11
; Sequence 11, Application US/09786240
; Patent No. 6558935
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GAUGLER, Karl J.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592 PCT
; CURRENT APPLICATION NUMBER: US/09/786, 240
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 60/1133,642
; NUMBER OF SEQ ID NOS: 33
; SEQ ID NO 11
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6558935 2617407CD1
; US-09-786-240-11

Query Match 47.7%; Score 981; DB 4; Length 386;
Best Local Similarity 54.8%; Pred. No. 1.5e-95;
Matches 215; Conservative 42; Mismatches 110; Indels 30; Gaps 9;

Qy 14 LLLAQTFULLFLVSRP-----GPSSPAGGEARHVYLVLSSWRSGSFVGQL 59
Db 2 LLPKMKLFLFLVSMALLFFMYSNHSNISLMSMKAQPERMHLVLSWRSGSFVGQL 61

Qy 119 DLFOAVASRALCSPACSAFPRGAISSCAVKPLICAROSTTLEACRSYSHVILKEYRF 177
Db 122 SLFQWENSRALCSAPACDIPQDEIIPRAHCRLJCSCQPFVEKACSYSHVILKEYR 180

Qy 178 FPNLQVLYPLISDPALNLRIVHLVRDPRAVRSRETOAKALARDNGIVLTGNTW-BAD 236
Db 181 FFNQIQLSYPLIKDPSLNHLHVLVRDPRAVFRSRETKGDLMDSRIVNGQHEQKLKED 240

Qy 237 PGURVYREVCRSHVRIAEAATLKP_PPELRGRYLVRFEDLAREPLAETRALLYFTGLSL 295
Db 241 QPYVMQVICQSQLEIYK-TIQSLPKALQERYLVRLVRDALARPAVQTSRMYBEGGLEF 298

Qy 296 TPQLEWAHNITHGSGP GARREAFKTSSENALNSQAMRHFAKIRRVQELCAGALQQL 355
Db 299 LPHQOTWVNITRGKGMG-DHAFTNARDALNSQAMRWSLPEKVSRLQACGDMNL 356

Qy 358 FNQLYPLISDPALNLRIVHLVRDPRAVRSRETOAKALARDNGIVLTGNTW-EADP 237
Db 359 FNLSQYPLIKDPSLNHLTHVLRDPRAVRSRETKGDMLIDSRYINGQHEQKLKKEDQ 241

US-09-190-911-1
; Sequence 1, Application US/09190911
; Patent No. 6365365
; GENERAL INFORMATION:
; APPLICANT: Bisterup, Annette
; APPLICANT: Rosén, Stefan
; APPLICANT: Hemmerich, Stefan
; APPLICANT: Tandemann, Kirsten
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE -3
; FILE REFERENCE: 6510-107C1P
; CURRENT APPLICATION NUMBER: US/09/190,911
; CURRENT FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/045,284
; EARLIER FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSSQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 386
; TYPE: PRT
; ORGANISM: H. sapiens
; US-09-190-911-1

Query Match 49.6%; Score 1019; DB 3; Length 386;
Best Local Similarity 54.5%; Pred. No. 1.4e-99;
Matches 216; Conservative 45; Mismatches 107; Indels 28; Gaps 7;

Qy 14 LLLAQTFULLFLVSRP-----GPSSPAGGEARHVYLVLSSWRSGSFVGQL 59
Db 2 LLPKMKLFLFLVSMALLFFMYSNHSNISLMSMKAQPERMHLVLSWRSGSFVGQL 61

Qy 60 FNQHPDVYLMEPAWHWTLSQSGSAATLHMAYRDLVRSLVFLCDMDVDAYL_PWRNLIS 118
Db 62 FGQHPDVYLMEPAWHWTLSQSGSAATLHMAYRDLVRSLVFLCDMSVFDAYMGPPIQS 121

Qy 119 DLFOAVASRALCSPACSAFPRGAISSCAVKPLICAROSTTLEACRSYSHVILKEYRF 178
Db 122 SLFQWENSRALCSAPACDIPQDEIIPRAHCRLJCSCQPFVEKACSYSHVILKEYR 181

Qy 179 FNQLYPLISDPALNLRIVHLVRDPRAVRSRETOAKALARDNGIVLTGNTW-EADP 237
Db 180 FNLSQYPLIKDPSLNHLTHVLRDPRAVRSRETKGDMLIDSRYINGQHEQKLKKEDQ 241

```

; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Marumatsu, Takashi
; TITLE OF INVENTION: DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM41.001AUS
; CURRENT APPLICATION NUMBER: US/09/471,867
; CURRENT FILING DATE: 1999-12-23
; EARLIER APPLICATION NUMBER: US/09/263,023
; EARLIER FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: JP 10-54007
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: JP 10-177844
; EARLIER FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 2
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-471-857-2

Query Match          32.2%; Score 663; DB 4; Length 483;
Best Local Similarity 39.5%; Pred. No. 1.4e-11;
Matches 146; Conservative 67; Mismatches 123; Indels 34; Gaps 9;

Qy          36  GGEARHVTLVLSWRSGSFSVQGLFNOHQDPDFYLMEPAWIVWTLSQGSATLNAVRDL 95
Db          113  GGDYRQLYVPTWRSSESSFFGELTQNPEFLTEPVWYWQKLXYPGDSVLQGARDM 172
Qy          96  VRSVFLCDMDYFDAYLP--KRRNLSDSL--RQWAVSRALCSPPACSAFPQGAIS--SEAV 148
Db          173  LSAYLYRCDLSVPLQYSPAGSSGRNLITLGFAATGNKVKVQVSPCPAYKEVVGJLDDRV 232
Qy          149  CRLPCARQSFITLAREQSYLQVYIPLSLDSDPANLRVHLLVDRPVAL 208
Db          233  CKK-CPQRLLARFEERKRYTVVIGVRYEDVAVLAFLPKDALKVTHLVRPRAVA 291
Qy          209  RSRQQATAKALARDN-----GIVLGT--NGTWVEAD-PGLRUVREV 245
Db          292  SSRIRSRHGLKRESLQVRSRDPRAHMPFLEAAGHKLGAKKEGMCGPDRYHALGAMEV 351
Qy          246  CRSHVRIIAEATLKKPPFLRGYRLVRFEDLAREPLAEIRALYAFTGLSLTPOLAWIHIN 305
Db          352  CNSMAKLTQA-LQPPDWLQCHLVYRVDLGVDPVTKLRRVDFGLVSPMEMQFALN 410
Qy          306  ITHGSGPQARRAEEFKTSRNLNVSAWHRALPFKIRYRQELCAGALQLGIRIVYSED 365
Db          411  MTSGSGSSSK - PFWVSARNATQAAANAWRTALTFQOIKQVEEFCTYQPMAVIGYERYNSPE 468
Qy          366  EQRNLLADLV 375
Db          469  EYRDLSKTLL 478

RESULT 9
US-09-949-016-9449
; Sequence 9449, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO/01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ For Windows Version 4.0
; SEQ ID NO: 9449
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-9449

Query Match Score 657.5; DB 4; Length 608;
Best Local Similarity 38.2%; Pred. No. 7.6e-61;
Matches 147; Conservative 70; Mismatches 125; Indels 43; Gaps 10;

Qy 30 GPSSPAG-----GEARHVLVLSWRSGSSEPVGOLFNOHPDVYFYLMEPAWHWNTL 80
Db 223 GVAAPPNGTRGNTGVDKRLQYVFTWRSGSSEFFGELFNQNPPEVFPLYEPWHWYQKL 282

Qy 81 SGQSAATHLMAVRDVLRSVFLCDMDVFDAYLP---WRNLSDL--FQWAVSRLCSPAC 135
Db 283 YPGDAVSLOGAARDMLSLAVALRCDSVFLQYSPAGSGGRNLTLIGFAGATNKVCSPLC 342

Qy 136 SAFFRGAIS--SEAVCKPLCAROSFTLAREACRSYSHVYLKEYRFNFNOLQVLYPLSDPAL 193
Db 343 PAYKEWVGLDDRVCK-KPPQBLARFEECRKYRTLVIGVRFVEDAVLAPLRLDPAL 401

Qy 194 NLRIVHLVRDPRAVLRSREQTAKLARD-----GIVLTG--NGT 231
Db 402 DLKVTHLVRDPRAVASSRTRSRHQLIREELQVYRSRDPRAHMPFLBAAGHKLGAKEGV 461

Qy 232 WVEAD-PGLRVREVCRSHVRIAFAATLKPFPFLRGRVLVRFEDLAPBLAERALYAF 290
Db 462 GGPDYHALQAMEVICNSMKTLOTA-LQPDWLNQGHYLVRVRLVGDGVPKTLRRTDF 520

Qy 291 TGLSLTPOLEAWHNITHGSGPGRARAEFKTSSENALANYSQMRHALPEAKTRVOLCA 350
Db 521 VGLVSPENBQFANMTGSSSK--PFVVSARNATQANANRALTFOQIKQVEEFCY 578

Qy 351 GALOLQYRPPVYSEDEQRNLALDV 375
Db 579 QPMAVLGYVERVNSPBEVDSLTKUL 603

RESULT 10
US-09-263-023-4 Application US/09263023
; Patenr No: 6037159
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA1_001AUS
; CURRENT APPLICATION NUMBER: US/09/471, 867
; CURRENT FILING DATE: 1999-12-23
; EARLIER APPLICATION NUMBER: US 09/263 , 023
; EARLIER FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: JP 10-54007
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: JP 10-177844
; EARLIER FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO: 4
; NUMBER OF SEQ ID NOS: 10
; FILE REFERENCE: TOYAMA1_001AUS
; CURRENT APPLICATION NUMBER: US/09/263, 023
; CURRENT FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: JP 10-54007
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: JP 10-177844
; EARLIER FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ For Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-263-023-4

Query Match Score 653.5; DB 3; Length 484;
Best Local Similarity 38.8%; Pred. No. 1.4e-60;
Matches 145; Conservative 70; Mismatches 124; Indels 35; Gaps 10;

Qy 33 SPAG-GEARHVLVLSWRSGSSEPVGOLFNOHPDVYFYLMEPAWHWNTL 91
Db 110 APPGVGDRHWMYVFTWRSGSSEFFGELFNQNPPEVFPLYEPWHWYQKL 169

Qy 92 VRDLVRSVFLCDMDVFDAYLP---WRNLSDL--FQWAVSRLCSPAC 91
Db 170 ARDMLSLAVALRCDSVFLQYSPAGSGGRNLTLIGFAGATNKVCSPLC 169

Qy 145 SEAVCKPLCAROSFTLAREACRSYSHVYLKEYRFNFNOLQVLYPLSDPALNRLVHLVRDP 204
Db 230 DDRVCK-KPPQBLARFEECRKYRTLVIGVRFVEDAVLAPLRLDPALDKVTHLVRDP 288

Qy 205 RAVLRSREQTAKLARD-----GIVLTG--INGTWEAD-PGLRV 241
Db 289 RAVASSRTRSRHQLIREELQVYRSRDPRAHMPFLBAAGHKLGAKEGVGPDPYHALGA 348

Qy 242 VREVCRSVHVRIAFAATLKPPFLRGRVLVRFEDLAPBLAERALYAF 301
Db 349 MEVICNSMKTLOTA-LQPDWLNQGHYLVRVRLVGDGVPKTLRRTDFVGLVSPENBQ 407

Qy 302 WHNITHGSGPGRARAEFKTSSENALANYSQMRHALPEAKTRVOLCA 361
Db 408 FALNMSSGSSSK--PFVVSARNATQANANRALTFOQIKQVEEFCYQPMAVLGYERYV 465

Qy 362 YSEDEQRNLALDV 375
Db 466 NSPPEVDSLTKUL 479

RESULT 11
US-09-471-867-4
; Sequence 4, Application US/09471867
; Patent No. 645289
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadomatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA1_001AUS
; CURRENT APPLICATION NUMBER: US/09/471, 867
; CURRENT FILING DATE: 1999-12-23
; EARLIER APPLICATION NUMBER: US 09/263 , 023
; EARLIER FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: JP 10-54007
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: JP 10-177844
; EARLIER FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO: 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-471-867-4

Query Match Score 653.5; DB 4; Length 484;
Best Local Similarity 38.8%; Pred. No. 1.4e-60;
Matches 145; Conservative 70; Mismatches 124; Indels 35; Gaps 10;

Qy 33 SPAG-GEARHVLVLSWRSGSSEPVGOLFNOHPDVYFYLMEPAWHWNTL 91
Db 110 APPGVGDRHWMYVFTWRSGSSEFFGELFNQNPPEVFPLYEPWHWYQKL 169

Qy 92 VRDLVRSVFLCDMDVFDAYLP---WRNLSDL--FQWAVSRLCSPAC 91
Db 170 ARDMLSLAVALRCDSVFLQYSPAGSGGRNLTLIGFAGATNKVCSPLC 169

Qy 145 SEAVCKPLCAROSFTLAREACRSYSHVYLKEYRFNFNOLQVLYPLSDPALNRLVHLVRDP 204
Db 230 DDRVCK-KPPQBLARFEECRKYRTLVIGVRFVEDAVLAPLRLDPALDKVTHLVRDP 288

Db 230 DDRVCKK-CPPQRLARFEECRKYRTLVIGYRVEDVAVLAPILRDLKVTHLVRDP 288
 Qy 205 RAYLRSREOTAKALARDN-----GIVLGT--NGTWBEAD-PGLRV 241
 Db 289 RAVASSRIRSRRHQLIRESLQQVRSRDPRAHMPFLEAGHKUGAKKEGVGGPADDYHALGA 348
 Qy 242 VREVCRSVHRVIAFAATLKKPPLFLGRYRVLRFEDLAEPPLAETRALYAFTGLSLTPOLEA 301
 Qy 349 MEVICNSMAKLTQTA-LQPPDWIQLGHYLVRVYEDLVDGPVKTLRRLVYDFGVGLVSPMEQ 407
 Db 302 WHNITHGSGPGRAREAPTKSENNALNVSQAWHLPKAIRVQELCAGALQLGYGRPV 361
 Qy 408 FALNMTCSSSSSK- PTVVSARNATQANANWARTLFQIQKVEEFCYQPMAVLGYERV 465
 Qy 362 YSEDEORNLAIDLV 375
 Db 466 NSBEEVKDLSTKL 479

RESULT 12
 US-09-949-016-6471
 Sequence 6471, Application US/094949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL-001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SEQ ID NO: 6471
 LENGTH: 531
 TYPE: PRT
 ORGANISM: Human
 us-09-949-016-6471

Query Match 31.8%; Score 653.5; DB 4; Length 531;
 Best Local Similarity 38.8%; Prod. No. 1.6e-60;
 Matches 145; Conservative 70; Mismatches 124; Indels 35; Gaps 10;

Db 33 SPAG-GEARVHVLSSNRSGSSFVGQLFNQHPDVFLMPEPAWHWHTLTSQSGAAATHMA 91
 Db 157 APEVGDRKRHHVYFTTWSRGSSFFGFLFNQNPVEVFFFYEPWVHWYKLYPEDAVS1QGA 216
 Qy 92 VRDLVRSVFLCDMDVFDAYLP---WRNLJSDL--PQWAVSRALCSPACASPRGAI---
 Qy 217 ARDMLSYRCDSVFLQLYSPAGSGGRNLTTLGIFGAATNKVCCSPCPARKEVVGLV 276
 Qy 145 SEAVCKPLCARQSFETLAREACRSYSHVYLKEYRFENLQVLYPLSDPALNLRVHLVRDP 204
 Db 277 DDRVCKK-CPPQRLARFEECRKYRTLVIGYRVEDVAVLAPILRDLKVTHLVRDP 335
 Qy 205 RAYLRSREOTAKALARDN-----GIVLGT--NGTWBEAD-PGLRV 241
 Db 336 RAVASSRIRSRRHQLIRESLQQVRSRDPRAHMPFLEAGHKUGAKKEGVGGPADDYHALGA 395
 Db 242 VREVCRSVHRVIAFAATLKKPPLFLGRYRVLRFEDLAEPPLAETRALYAFTGLSLTPOLEA 301
 Qy 396 MEVICNSMAKLTQTA-LQPPDWIQLGHYLVRVYEDLVDGPVKTLRRLVYDFGVGLVSPMEQ 454
 Qy 302 WHNITHGSGPGRAREAPTKSENNALNVSQAWHLPKAIRVQELCAGALQLGYGRPV 361
 Qy 455 FALNMTCSSSSSK- PTVVSARNATQANANWARTLFQIQKVEEFCYQPMAVLGYERV 512
 Db 362 YSEDEORNLAIDLV 375

Db 513 NSBEEVKDLSTKL 526

RESULT 13
 US-08-899-514-2
 Sequence 2, Application US/08899514
 Patent No. 5910581
 GENERAL INFORMATION:
 APPLICANT: HABUCHI, OSAMI
 APPLICANT: FUKUDA, MASAKAZU
 TITLE OF INVENTION: POLYPEPTIDE OF GLYCOSAMINOGLYCAN
 TITLE OF INVENTION: SULFOTRANSFERASE ORIGINATING FROM HUMAN AND DNA CODING
 TITLE OF INVENTION: FOR THE POLYPEPTIDE
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: KNOBBE, MARTENS, OLSON & BEAR, LLP
 STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
 CITY: NEWPORT BEACH
 STATE: CALIFORNIA
 COUNTRY: US
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/899,514
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 ATTORNEY/AGENT INFORMATION:
 NAME: DANIEL E ALTMAN
 REGISTRATION NUMBER: 34,115
 REFERENCE/DOCKET NUMBER: TOYAM21.001AUS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 714 760 0404
 TELEFAX: 714 760 9502
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 479
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 us-08-899-514-2

Query Match 28.0%; Score 576.5; DB 2;
 Best Local Similarity 37.2%; Pred. No. 2.2e-52;
 Matches 136; Conservative 58; Mismatches 141; Indels 31; Gaps 8;

Qy 87 TLHMAVDLVLRSVFLCDMDVFDAY--LPWRLNSLDFOWAYSRALCSPACASAFPGAI 143
 Db 183 GSALVYDVLQFLCLQVLYHFITLQFRGSSRSLSQBDPVCPTPVKVF 242
 Qy 31 PSSPAGGEARVHVLVLSWSRGSSFSYGVOLFNFQHPDVFLMPEPAWHWHTLTS---QGSAA 86
 Db 123 PPRPAVAGPRRHVLLMATTRTSSFVGFBNQGNITYLFEPFLWHRTVSFEPGGANAA 182

Qy 144 SSEAVCKPLCARQSFETLAREACRSYSHVYLKEYRFENLQVLYPLSDPALNLRVHLVRDP 203
 Db 243 EYHCKNRRQRCGPINVTAAECCRKEHMLAKVRIQLEFLQPLAEPRFLRVIQVLRD 302
 Db 180 PRAVLRSLREQTAKALARDNGTIVLGTNTGTVW---EADPGJRV---VREVCRSHYRIA 253
 Db 303 PRAVLAHS---MVAPA---GKTYTKWKLDDQGDLREEVQLRGNCES-IRLS 351
 Qy 204 PRAVLRSLREQTAKALARDNGTIVLGTNTGTVW---EADPGJRV---VREVCRSHYRIA 253
 Db 352 AEGLRQPAWLGRYMLVRYEDVARGLQKAREMPPAGIPLTPQvdwiQkntQAAHDG 411

RESULT 14

Qy 314 ARREAFKTSRNALNVSQMRHALPFAKIRVQECLCAGALQLIGYRPVYSEDEQRNLALD 373
Db 412 S---GIVSTQKNSSEQFERWSPNPKLQQVQAPCGPAMRLPCKLARDAAALTNSVS 468

Qy 374 LVLPRG 379
Db 469 LLEBERG 474

Query Match 27.0%; Score 555.5; DB 3; Length 411;
Best Local Similarity 36.7%; Pred. No. 2.9e-50;
Matches 131; Conservative 56; Mismatches 153; Indels 17; Gaps 8;

Qy 40 RVHVLVLISSRSGSSFVGQFLNQHPDVFEVLMEPAWHWTI---SOGSA---ATLHMA 91
Db 60 KTHILILATTRSGSSFVGQFLNQHLDVFLFEPYLHVQNTLIPRFTQGRSPDRRVMLGA 119
Qy 92 PWRNLSSDLFQWAVSALCSPPACAS-AFPGAISSA 147
Db 120 SRDLRSLYCDLIFPLENYIKPPPNHTTDRIFRRGASRVLCSPVCDPPGPADLVLEG 179
Qy 148 VCKPLCARQSFITLAREACRSYSHVYLKEYRFENLQVLYPLSDPALNLRVHLVYRDPRAV 207
Db 180 DCVRKGGLUNLTAEECRERSHVAIKTVRPEVNDRLAIVEDPRNLKIVQLVRDPRTI 239

Qy 208 LRSRQTAKLARDNGIVLCTGNTWVEADPGLVYREVCRSHVRIAETATLKPEPLRGR 267
Db 240 LASRSETFRDTYRLWLTGTRKEYNLD--VTQLTTVCEDFNSNVSTGLMR-PPWLKGK 296
Qy 268 YRLVRFEDLAREPLAEIAYLAFTGSLSLTQLEWIHNTHFGSGPGRARBEAKTSSRNAL 327
Db 297 YMLVRYEDLNRNPMMKTEEYFLGJPLDHSVARMIQNNTGRD-PTLGKHKYGT--VRNSA 354

Qy 328 NVSQMRHALPFAKIRVQECLCAGALQLIGYRPVYSEDEQRNLALDLYLVRGLNGFT 384
Db 355 ATAEMRFRFLSYDIVAFANQACQVQLAQZKIAASEEELKNPSVSLVERDFRPFS 411

RESULT 15

US-08-655-878-2

Sequence 2, Application US/086555878

Patent No. 6227713

GENERAL INFORMATION:

APPLICANT: FUKUTA, MASAKAZU
APPLICANT: HABUCHI, OSAMI
TITLE OF INVENTION: DNA CODING FOR SULFOTRANSFERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE:
STREET:
CITY:
STATE:
COUNTRY:

; ZIP: COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: FC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/655,878
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-655-878-2

Query Match 25.9%; Score 532.5; DB 2; Length 458;
Best Local Similarity 34.6%; Pred. No. 9.7e-48;
Matches 134; Conservative 63; Mismatches 157; Indels 33; Gaps 10;

Qy 5 RVSSTATALLAQTFLLFLFLYSPRGSSPAGBARYRVLVLSWRSSSFVQQLFNQHP 64
Db 80 RQDLDTFSQSLRTRLRNTLQLAGEGLAAP--EPRHVLVLLATTRCSSFPVGEFPNQOG 136
Qy 65 DVFYLMPEAHWY--WTTLSOGSRSATLIMAV--RDLVRSVELCDMDVFDAYL---PWRNRL 117
Db 137 NIVYLFEPWHIBRTVTPEGGANASALVYDVLQQLLCLXLYLESFISPAPEHIL 196
Qy 118 SDLQWAVSRLCSPPACASAPRGAISSAECVKPLCAROSFTLAREACRSYSHVVLKEVR 177
Db 197 AAFLRGGSSHSLCCEPVCTPSLKKVFBXCHCNTRCGPLNITLAEEACRKHOHMLAKTVR 256
Qy 178 FFNLQVLYPLLSDDPDLNRLIVHLVDRPAVLRREREQATAKALARDNGIVLGTNTW---V 233
Db 257 TROLEFLQPLAEDPDLRDIQLVDRPAVLSSRMVAFS-----GKYESWKWMAA 306
Qy 234 EADPGLRV----VREYCRSHVRIAETATLKPEPLRGRFLRFLDAREPLAEIRALY 288
Db 307 EGRAAPLOEDEVQLRGNCES-IRLSAEGLQRPMWLRGRMLRYEVAPAPRKALEY 365
Qy 289 AFTGLSLPQLEWMHNITHGSSPGRAREAFTSSRNALNVSQMRHALPFAKIRVQEL 348
Db 366 RFAGIHPIPQVEEWIRANT--QAPQDSNGTYST-QKNSSEQFEKWRSTIPFLAQVvQDPA 422
Qy 349 CAGALQIIGYRPVYSEDEQRNLALDLY 375
Db 423 CEPMARLFGYKCLASSAQLTNRSSL 449

Search completed: June 23, 2005, 08:52:14
Job time : 15.8883 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM Protein - Protein search, using SW model

Run on: June 23, 2005, 08:49:14 ; Search time 49.935 Seconds

(without alignments)

3041.886 Million cell updates/sec

Title: US-10-697-828-13
 Perfect score: 2056
 Sequence: 1 MWLPRVSSTAVTLLAQTF.....LPRGLNGFTWASSTASHPRN 395Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 10%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cg92_6/ptodata/1/pubpa/us07_pubcomb.pep:*

2: /cg92_6/ptodata/1/pubpa/us05_pub:*

3: /cg92_6/ptodata/1/pubpa/us06_pub:*

4: /cg92_6/ptodata/1/pubpa/us05_pubcomb.pep:*

5: /cg92_6/ptodata/1/pubpa/us07_pub:*

6: /cg92_6/ptodata/1/pubpa/pctus_pubcomb.pep:*

7: /cg92_6/ptodata/1/pubpa/us08_pub:*

8: /cg92_6/ptodata/1/pubpa/us09_pubcomb.pep:*

9: /cg92_6/ptodata/1/pubpa/us09_pubcomb.pep:*

10: /cg92_6/ptodata/1/pubpa/us09_pubcomb.pep:*

11: /cn2_6/ptodata/1/pubpa/us09_pubcomb.pep:*

12: /cn2_6/ptodata/1/pubpa/us09_pub:*

13: /cn2_6/ptodata/1/pubpa/us10_pubcomb.pep:*

14: /cn2_6/ptodata/1/pubpa/us10_pubcomb.pep:*

15: /cn2_6/ptodata/1/pubpa/us10_pubcomb.pep:*

16: /cn2_6/ptodata/1/pubpa/us10_pubcomb.pep:*

17: /cn2_6/ptodata/1/pubpa/us10_pubcomb.pep:*

18: /cn2_6/ptodata/1/pubpa/us10_pub:*

19: /cn2_6/ptodata/1/pubpa/us11_pubcomb.pep:*

20: /cn2_6/ptodata/1/pubpa/us11_pub:*

21: /cn2_6/ptodata/1/pubpa/us09_pub:*

22: /cn2_6/ptodata/1/pubpa/us06_pubcomb.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|--|
| 1 | 2056 | 100.0 | 395 | 9 US-09-927-502-2 Sequence 2, Appli |
| 2 | 2056 | 100.0 | 395 | 15 US-10-258-080-5 Sequence 5, Appli |
| 3 | 2056 | 100.0 | 395 | 16 US-10-648-593-159 Sequence 6, Appli |
| 4 | 2056 | 100.0 | 395 | 16 US-10-697-828-13 Sequence 7, Appli |
| 5 | 1934 | 94.1 | 395 | 9 US-09-927-002-3 Sequence 13, Appli |
| 6 | 1729.5 | 84.1 | 390 | 9 US-09-127-602-4 Sequence 1, Appli |
| 7 | 1729.5 | 84.1 | 390 | 16 US-10-697-828-8 Sequence 3, Appli |
| 8 | 1708 | 83.1 | 395 | 16 US-10-697-828-7 Sequence 4, Appli |
| 9 | 1708 | 83.1 | 395 | 16 US-10-841-107-8 Sequence 5, Appli |
| 10 | 1694 | 82.4 | 418 | 9 US-09-927-502-5 Sequence 6, Appli |
| 11 | 1538 | 74.8 | 394 | 16 US-10-408-765A-2211, Ap |

US-09-927-602-2

US-09-927-602-3

US-09-927-602-4

US-09-927-602-5

US-09-927-602-6

US-09-927-602-7

US-09-927-602-8

US-09-927-602-9

US-09-927-602-11

US-09-927-602-12

US-09-927-602-13

US-09-927-602-14

US-09-927-602-15

US-09-927-602-16

US-09-927-602-17

US-09-927-602-18

US-09-927-602-19

US-09-927-602-20

US-09-927-602-21

US-09-927-602-22

US-09-927-602-23

US-09-927-602-24

US-09-927-602-25

US-09-927-602-26

US-09-927-602-27

US-09-927-602-28

US-09-927-602-29

US-09-927-602-30

US-09-927-602-31

US-09-927-602-32

US-09-927-602-33

US-09-927-602-34

US-09-927-602-35

US-09-927-602-36

US-09-927-602-37

US-09-927-602-38

US-09-927-602-39

US-09-927-602-40

US-09-927-602-41

US-09-927-602-42

US-09-927-602-43

US-09-927-602-44

US-09-927-602-45

US-09-927-602-46

US-09-927-602-47

US-09-927-602-48

US-09-927-602-49

US-09-927-602-50

US-09-927-602-51

US-09-927-602-52

US-09-927-602-53

US-09-927-602-54

US-09-927-602-55

US-09-927-602-56

US-09-927-602-57

US-09-927-602-58

US-09-927-602-59

US-09-927-602-60

US-09-927-602-61

US-09-927-602-62

US-09-927-602-63

US-09-927-602-64

US-09-927-602-65

US-09-927-602-66

US-09-927-602-67

US-09-927-602-68

US-09-927-602-69

US-09-927-602-70

US-09-927-602-71

US-09-927-602-72

US-09-927-602-73

US-09-927-602-74

US-09-927-602-75

US-09-927-602-76

US-09-927-602-77

US-09-927-602-78

US-09-927-602-79

US-09-927-602-80

US-09-927-602-81

US-09-927-602-82

US-09-927-602-83

US-09-927-602-84

US-09-927-602-85

US-09-927-602-86

US-09-927-602-87

US-09-927-602-88

US-09-927-602-89

US-09-927-602-90

US-09-927-602-91

US-09-927-602-92

US-09-927-602-93

US-09-927-602-94

US-09-927-602-95

US-09-927-602-96

US-09-927-602-97

US-09-927-602-98

US-09-927-602-99

US-09-927-602-100

US-09-927-602-101

US-09-927-602-102

US-09-927-602-103

US-09-927-602-104

US-09-927-602-105

US-09-927-602-106

US-09-927-602-107

US-09-927-602-108

US-09-927-602-109

US-09-927-602-110

US-09-927-602-111

US-09-927-602-112

US-09-927-602-113

US-09-927-602-114

US-09-927-602-115

US-09-927-602-116

US-09-927-602-117

US-09-927-602-118

US-09-927-602-119

US-09-927-602-120

US-09-927-602-121

US-09-927-602-122

US-09-927-602-123

US-09-927-602-124

US-09-927-602-125

US-09-927-602-126

US-09-927-602-127

US-09-927-602-128

US-09-927-602-129

US-09-927-602-130

US-09-927-602-131

US-09-927-602-132

US-09-927-602-133

US-09-927-602-134

US-09-927-602-135

US-09-927-602-136

US-09-927-602-137

US-09-927-602-138

US-09-927-602-139

US-09-927-602-140

US-09-927-602-141

US-09-927-602-142

US-09-927-602-143

US-09-927-602-144

US-09-927-602-145

US-09-927-602-146

US-09-927-602-147

US-09-927-602-148

US-09-927-602-149

US-09-927-602-150

US-09-927-602-151

US-09-927-602-152

US-09-927-602-153

US-09-927-602-154

US-09-927-602-155

US-09-927-602-156

US-09-927-602-157

US-09-927-602-158

US-09-927-602-159

US-09-927-602-160

US-09-927-602-161

US-09-927-602-162

US-09-927-602-163

US-09-927-602-164

US-09-927-602-165

US-09-927-602-166

US-09-927-602-167

US-09-927-602-168

US-09-927-602-169

US-09-927-602-170

US-09-927-602-171

US-09-927-602-172

US-09-927-602-173

US-09-927-602-174

US-09-927-602-175

US-09-927-602-176

US-09-927-602-177

US-09-927-602-178

US-09-927-602-179

US-09-927-602-180

US-09-927-602-181

US-09-927-602-182

US-09-927-60

Query Match Similarity 100.0%; Score 2056; DB 15; Length 395;
 Best Local Similarity 100.0%; Pred. No. 4..6..198;
 Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWLPRVSTAVTALLAQTFLFLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 60
 Db 1 MWLPRVSTAVTALLAQTFLFLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 60

Qy 61 NQHPDVYLMERAMHVTLSQGSAATLMAVDLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 120
 Db 61 NQHPDVYLMERAMHVTLSQGSAATLMAVDLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 120

Qy 61 FQAVASRALCSPPACSAFPRAISSAVCPKLCARQSFTLAREACRSYSHVVLKEVRFFN 180
 Db 121 FQAVASRALCSPPACSAFPRAISSAVCPKLCARQSFTLAREACRSYSHVVLKEVRFFN 180

Qy 121 FQAVASRALCSPPACSAFPRAISSAVCPKLCARQSFTLAREACRSYSHVVLKEVRFFN 180
 Db 121 FQAVASRALCSPPACSAFPRAISSAVCPKLCARQSFTLAREACRSYSHVVLKEVRFFN 180

Qy 181 LQVLYPLISDPALNLRIVHLVRDPAVLRSRECTAKALARNDGIVLGNTGTVWEADPGLR 240
 Db 181 LQVLYPLISDPALNLRIVHLVRDPAVLRSRECTAKALARNDGIVLGNTGTVWEADPGLR 240

Qy 241 VREVCRSHVRIAATAILKPPPLERGRYPLVRFEDLAPELARIALYAFGLSLTQLE 300
 Db 241 VREVCRSHVRIAATAILKPPPLERGRYPLVRFEDLAPELARIALYAFGLSLTQLE 300

Qy 301 AWINITHSSGPGRARKEATPSNNALAVSQMRHALPFKAIRRQELCAGALQLGYRP 360
 Db 301 AWINITHSSGPGRARKEATPSNNALAVSQMRHALPFKAIRRQELCAGALQLGYRP 360

Qy 361 VYSEDEQRNLALDLVLPRLGNGFTWASSTA SHPRN 395
 Db 361 VYSEDEQRNLALDLVLPRLGNGFTWASSTA SHPRN 395

Qy 241 VYREVCSHYRIAETALKPPLERGRYPLVRFEDLAPELARIALYATGLSLTQLE 300
 Db 241 VYREVCSHYRIAETALKPPLERGRYPLVRFEDLAPELARIALYATGLSLTQLE 300

Qy 301 AWINITHSSGPGRARKEATPSNNALAVSQMRHALPFKAIRRQELCAGALQLGYRP 360
 Db 301 AWINITHSSGPGRARKEATPSNNALAVSQMRHALPFKAIRRQELCAGALQLGYRP 360

Qy 361 VYSEDEQRNLALDLVLPRLGNGFTWASSTA SHPRN 395
 Db 361 VYSEDEQRNLALDLVLPRLGNGFTWASSTA SHPRN 395

RESULT 3
 US-10-648-593-159
 ; Sequence 159 Application US/10648593
 ; Publication No. US20041002125A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Incyte Genomics, Inc..
 ; POLICKY, Jennifer L.
 ; HAPALIA, April J.A.
 ; BURFORD, Neil
 ; RING, Huijun Z.
 ; LAL, Preet G.
 ; TRIBOLLEY, Catherine M.
 ; YAO, Monique G.
 ; YUE, Henry
 ; TANG, Y. Tom
 ; ARVIZU, Chandra S.
 ; DAS, Debopriya
 ; SANJANWALA, Madhusudan M.
 ; GANDHI, Ameeta R.
 ; REDDY, Roopa M.
 ; KHAN, Farrah A.
 ; BAUGHN, Mariyah R.
 ; RANKUMAR, Jayalaxmi
 ; GRIFFIN, Jennifer A.
 ; AU-YOUNG, Janice K.
 ; PRIORITY INVENTION: DRUG METABOLIZING ENZYMES
 ; FILE REFERENCE: PI-0070 USN
 ; CURRENT APPLICATION NUMBER: US/10/258, 080
 ; CURRENT FILING DATE: 2002-10-15
 ; PRIOR APPLICATION NUMBER: US 60/203, 509
 ; PRIOR FILING DATE: 2000-05-11
 ; PRIOR APPLICATION NUMBER: US 60/202, 234
 ; PRIOR FILING DATE: 2000-05-05
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 5
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20040029125A1 7472777CD1
 ; NAME/KEY:
 ; US-10-648-593-159

Query Match Similarity 100.0%; Score 2056; DB 16; Length 395;
 Best Local Similarity 100.0%; Pred. No. 4..6..198;
 Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWLPRVSTAVTALLAQTFLFLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 60
 Db 1 MWLPRVSTAVTALLAQTFLFLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 60

Qy 61 NQHPDVYLMERAMHVTLSQGSAATLMAVDLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 120
 Db 61 NQHPDVYLMERAMHVTLSQGSAATLMAVDLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 120

Qy 121 FQAVASRALCSPPACSAFPRAISSAVCPKLCARQSFTLAREACRSYSHVVLKEVRFFN 180
 Db 121 FQAVASRALCSPPACSAFPRAISSAVCPKLCARQSFTLAREACRSYSHVVLKEVRFFN 180

Qy 181 LQVLYPLISDPALNLRIVHLVRDPAVLRSRECTAKALARNDGIVLGNTGTVWEADPGLR 240

Db 181 LQVLYPLLSDPALNLRIVHLVRDPRAVLRSREQTAKALARDNGIVLGTNGTWADPGLR 240
 Qy 241 VREYCRSHRIAEATLKEPPFLGRYRIVRFLRFDLAREPLAEIRALYAFGLSLTPQE 300
 Db 241 VREYCRSHRIAEATLKEPPFLGRYRIVRFLRFDLAREPLAEIRALYAFGLSLTPQE 300
 ; Sequence 3, Application US/09927602-3
 ; Patent No. US20020061562A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Michiko N.
 ; INVENTION: Akama, Tomoya O.
 ; TITLE OF INVENTION: Methods of Treating Macular Corneal
 ; FILE REFERENCE: P-LJ 4852
 ; CURRENT APPLICATION NUMBER: US/09/927-602
 ; CURRENT FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: US 09/638, 211
 ; PRIOR FILING DATE: 2000-08-11
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: Fast-SEQ For Windows Version 4.0
 ; SEQ ID NO: 3
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetic construct
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)..(395)
 ; OTHER INFORMATION: xaa = any amino acid
 ; US-09-927-602-3

RESULT 4

US-10-697-828-13
 ; Sequence 13, Application US/1067828
 ; Publication No. US2004010185546A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosenthal, Steven
 ; INVENTION: Novel Glycosyl sulfotransferases GST-4alpha, & GST-6
 ; FILE REFERENCE: UCAL-138D1V
 ; CURRENT APPLICATION NUMBER: US/10/697, 828
 ; CURRENT FILING DATE: 2003-10-29
 ; PRIOR APPLICATION NUMBER: 09/533, 828
 ; PRIOR FILING DATE: 2000-06-13
 ; PRIOR APPLICATION NUMBER: 60/144, 694
 ; PRIOR FILING DATE: 1999-07-20
 ; NUMBER OF SEQ ID NOS: 23
 ; SEQ ID NO: 13
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; US-10-697-828-13

Query Match 94.1%; Score 1934; DB 9; Length 395;
 Best Local Similarity 94.4%; Pred. No. 9_1e-186;
 Matches 371; Conservative 3; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MWLPRVSSTAVTALLAQTFLFLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 60
 Db 1 MWLPRFSSXTVXLLAQITXLFLVSRGPSSPAGGEARHVLVLSWRSGSSFVGQLF 60
 Qy 61 NQHPDVYLMEPAWHVVNTLQSQAATLHMAYRDLVRSVFLCDMDVDAYLPWRRNLSDL 120
 Db 61 SQHPDVFYLMEPAWHVVNTLQSQAATLHMAYRDLXRSVFLCDMDVDAYLPWRRNLSDL 120
 Qy 121 FQWASRALCSPPACASAPRGAISSEAVCKPLCARQSFTLARACRYSHVULKEVRFFN 180
 Db 121 FQWASRALCSPPACASAPRGSISSEVKPLCARQSFTLARACRYSHVULKEVRFFN 180
 Qy 181 LQVLYPLISDPAINRLTVHLVRDPRAVLRSREQTAKALARDNGIVLGTNGTWADPGLR 240
 Db 181 LQVLYPLISDPAINRLTVHLVRDPRAVLRSREQTAKALARDNGIVLGTNGTWADPGLR 240
 Qy 241 VREYCRSHVRIAATLKEPPFLGRYRIVRFLRFDLAREPLAEIRALYAFGLSLTPQE 300
 Db 241 VREYCRSHVRIAATLKEPPFLGRYRIVRFLRFDLAREPLAEIRALYAFGLSLTPQE 300
 Qy 301 AWIHNITGSGP GARREFAKFTSSRNALAVSQMPHALPPAKIRVQELCAGAQQLGGRP 360
 Db 301 AWIHNITGSGP GARREFAKFTSSRNALAVSQMPHALPPAKIRVQELCAGAQQLGGRP 360
 Qy 361 VYSEDEORNLAIDLVPRLGLNGPTWAASSASHP 393
 Db 361 VYSEDEQRDLXDLQDLPQGDXFWWAASSXXP 393
 ; Sequence 4, Application US/09927602-4
 ; Patent No. US20020061562A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Michiko N.
 ; INVENTION: Akama, Tomoya O.
 ; TITLE OF INVENTION: Methods of Treating Macular Corneal
 ; FILE REFERENCE: P-LJ 4852
 ; CURRENT APPLICATION NUMBER: US/09/927-602
 ; CURRENT FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: US 09/638, 211
 ; PRIOR FILING DATE: 2000-08-11

RESULT 5

Db 181 LQVLYPLLSDPALNLRIVHLVRDPRAVLRSREQTAKALARDNGIVLGTNGTWADPGLR 240
 Qy 241 VREYCRSHRIAEATLKEPPFLGRYRIVRFLRFDLAREPLAEIRALYAFGLSLTPQE 300
 Db 241 VREYCRSHRIAEATLKEPPFLGRYRIVRFLRFDLAREPLAEIRALYAFGLSLTPQE 300
 ; Sequence 5, Application US/09927602-5
 ; Patent No. US20020061562A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Michiko N.
 ; INVENTION: Akama, Tomoya O.
 ; TITLE OF INVENTION: Methods of Treating Macular Corneal
 ; FILE REFERENCE: P-LJ 4852
 ; CURRENT APPLICATION NUMBER: US/09/927-602
 ; CURRENT FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: US 09/638, 211
 ; PRIOR FILING DATE: 2000-08-11

60 FNOHPDVFYLMEPAWHWTLSQGSAAATHMARDLVRDVFDAYLPWRRNISD 119
 Qy ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 4
 ; LENGTH: 390
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-09-927-602-4

Query Match 84.1%; Score 1729.5; DB 9; Length 390;
 Best Local Similarity 85.8%; Bred. No. 3.6e-165;
 Matches 333; Conservative 17; Mismatches 37; Indels 1; Gaps 1;

Db 1 MWLPRVSSTAATLLAQ-TPLLFLYSPGPSSPAGGEARVHVLSSRSGSSFVGQL 59
 Db 1 MWLPRFESKTTVLLAQTTCCLFLISRPGPSSPAGGEARVHVLSSRSGSSFVGQL 60

Qy 60 FNOHPDVFYLMEPAWHWTLSQGSAAATHMARDLVRDVFDAYLPWRRNISD 119
 Db 61 FNOHPDVFYLMEPAWHWTLSQGSAAATHMARDLVRDVFDAYLPWRRNISD 120

Qy 120 LFOWAWSRALCSPPACSAFPFGAISSEAVCKPLCARQSFTLAREACRSYSHVVLKEVRFF 179
 Db 121 FPNWATSRALCSPPACSAFPFGTISKDVCXLTCTQPFSAREACRSYSHVVLKEVRFF 180

Qy 180 NLQVLYPLSDPAINLRIVHLYDPAVILRSREQTAKALARDNGIVLGNTWEDPGL 239
 Db 181 NLQVLYPLSDPAINLRIVHLYDPAVILRSREAGPILLARDNGIVLGNTKNYEADPHL 240

Qy 240 RVREVCRSHVRLAEATLKPPPLGRGRLVRFEDLREPLAEIRALYAFTGSLTLSQL 299
 Db 241 RLIREVCRSHVRLAEATLKPPPLGRGRLVRFEDLREPLAEIRALYAFTGSLTLSQL 300

Qy 300 EAWIHNTHGSGPGRAEFKTSRNALNVSQAWRHALPFAKIRRVQELCAGALQOLGYR 359
 Db 301 EAWIHNTHGSGIOPKPIEAFHTTSRNANVSQAWRHALPFTKILRVQEVCAQALQOLGYR 360

Db 360 PYVSEDEORNALDVLPRGLNGFTWAS 387
 Db 361 PYVADQRDLTDVLPRGDHFMSMAS 388

RESULT 8
 US-10-697-828-7
 ; Sequence 7, Application US/10697828
 ; Publication No. US20040185546A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, Steven
 ; INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-5
 ; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-5
 ; FILE REFERENCE: UCAL-138DIV
 ; CURRENT APPLICATION NUMBER: US/10/697.828
 ; CURRENT FILING DATE: 2003-10-29
 ; PRIOR APPLICATION NUMBER: 09/593,328
 ; PRIOR FILING DATE: 2000-06-13
 ; PRIOR APPLICATION NUMBER: 60/144,694
 ; PRIOR FILING DATE: 1999-07-20
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 7
 ; LENGTH: 395
 ; TYPE: PRT
 ; ORGANISM: mus musculus
 US-10-697-828-7

RESULT 7
 US-10-697-828-8
 ; Sequence 8, Application US/10697828
 ; Publication No. US20040185546A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen, Steven
 ; INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-6
 ; FILE REFERENCE: UCAL-138DIV
 ; CURRENT APPLICATION NUMBER: US/10/697,828
 ; CURRENT FILING DATE: 2003-10-29
 ; PRIOR APPLICATION NUMBER: 09/593,828
 ; PRIOR FILING DATE: 2000-06-13
 ; PRIOR APPLICATION NUMBER: 60/144,694
 ; PRIOR FILING DATE: 1999-07-20
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 8
 ; LENGTH: 390
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-697-828-8

Query Match 84.1%; Score 1729.5; DB 16; Length 390;
 Best Local Similarity 85.8%; Bred. No. 3.6e-165;
 Matches 333; Conservative 17; Mismatches 37; Indels 1; Gaps 1;

Db 1 MWLPRVSSTAATLLAQ-TPLLFLYSPGPSSPAGGEARVHVLSSRSGSSFVGQL 59
 Db 1 MWLPRFESKTTVLLAQTTCCLFLISRPGPSSPAGGEARVHVLSSRSGSSFVGQL 60

Qy 1 MWLPRVSSTAATLLAQ-TPLLFLYSPGPSSPAGGEARVHVLSSRSGSSFVGQL 59
 Db 1 MWLPRFESKTTVLLAQTTCCLFLISRPGPSSPAGGEARVHVLSSRSGSSFVGQL 60

Qy 121 FOWAWSRALCSPPACSAFPFGAISSEAVCKPLCARQSFTLAREACRSYSHVVLKEVRFF 180
 Db 121 FOWAWSRALCSPPCYCEAPRGNTSSEYCKPLCATRPGLAQACSSYSHVVLKEVRFF 180

Qy 181 LQVLYPLSDPAINLRIVHLYDPAVILRSREQTAKALARDNGIVLGNTWEDPGL 240
 Db 181 LQVLYPLSDPAINLRIVHLYDPAVILRSREAGPILLARDNGIVLGNTWEDPGL 240

Qy 241 VREVCRSHVRLAEATLKPPPLGRGRLVRFEDLREPLAEIRALYAFTGSLTLSQL 300
 Db 241 VREVCRSHVRLAEATLKPPPLGRGRLVRFEDLREPLAEIRALYAFTGSLTLSQL 300

RESULT 9

Qy 301 AWIHNITHGSGPGRAREAFKTSRNALNVSQAWRHALPFKKIRRVQELCAGALQQLGYRP 360
 Db 301 TWINNITHGSGPGRAREAFKTSRNALNVSQAWRHALPFKKIRRVQELCAGALQQLGYRS 360

Qy 361 VYSEDEQRNIALDVLVRGLNGFTWASSTASHPRN 395
 Db 361 VHSLEQRDLSIDLULPRGMDSFKMASSTEKOPES 395

RESULT 10

Qy 301 AWIHNITHGSGPGRAREAFKTSRNALNVSQAWRHALPFKKIRRVQELCAGALQQLGYRP 360
 Db 301 TWINNITHGSGPGRAREAFKTSRNALNVSQAWRHALPFKKIRRVQELCAGALQQLGYRS 360

Qy 361 VYSEDEQRNIALDVLVRGLNGFTWASSTASHPRN 395
 Db 361 VHSLEQRDLSIDLULPRGMDSFKMASSTEKOPES 395

; TITLE OF INVENTION: Methods of Treating Macular Corneal
 ; TITLE OF INVENTION: Diabetropy
 ; FILE REFERENCE: P-LJ 482
 ; CURRENT APPLICATION NUMBER: US/09/927,602
 ; CURRENT FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: US 09/638,211
 ; PRIOR FILING DATE: 2000-08-11
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 5
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-927-602-5

Query Match 82.4%; Score 1694.5; DB 9; Length 418;
 Best Local Similarity 82.8%; Pred. No. 1..3e-161;
 Matches 327; Conservative 24; Mismatches 43; Indels 1; Gaps 1;

Qy 1 MWLPRVSSATVATAQTEFLFLVSRGPSPAGEPARHVVLVLSWRSGSSFVGOLF 60
 Db 25 MRLPRESSTMISLMLMVTGILFLVSRQVPSSPAGLGERVHLVLSWRSGSSFVGQGF 84

Qy 61 NQHPDVFLMEPAWHWVTTLSQGSATLHMAVDLVRSLVFLCDMDVDAYLPWRRNLSDL 120
 Db 85 SQHPDVFLMEPAWHWVDTLSQGSAPALHMAVDLVRSLVFLCDMDVDAYLPWRRNLSDL 144

Qy 121 FQAVASRALCSPPACSAFPGASSEAVCPKLCAROSFTLARECRSYHVKLKEVRFFN 180
 Db 145 FQAVASRALCSPPVCFAFARGNSSEEVCPKLCATRPGLAQECSTSIVHVKLKEVRFFN 204

Qy 181 LQVLYPLSDPALNRLIHLVDRPAVLRSREQTAKALARDNGTVLGNTWYADPGLR 240
 Db 205 LQVLYPLSDPALNRLIHLVDRPAVLRSREQTAKALARDNGTVLGNTWYADPGLR 264

Qy 241 VVREVCRSVHRIAEBATLKPPFLGRYRLVRFEDLAREPLAEIRALYAFTGLSITPQLE 300
 Db 265 VVNEVCRSVHRIAEBALHKKPPFLQDRYRLVRFEDLARDPLTVIRELAFTGLGLTPQLO 323

Qy 301 AWIHNITHGSGPGRAREAFKTSRNALNVSQAWRHALPFKKIRRVQELCAGALQQLGYRP 360
 Db 324 TWIHNITHGSGPGRAREAFKTSRNALNVSQAWRHTLPPAKIRRVQELCGALQQLGYRS 383

Qy 361 VYSEDEQRNIALDVLVRGLNGFTWASSTASHPRN 395
 Db 384 VHSLEQRDLSIDLULPRGMDSFKMASSTEKOPES 418

RESULT 11

Qy 181 LQVLYPLSDPALNRLIHLVDRPAVLRSREQTAKALARDNGTVLGNTWYADPGLR 240
 Db 181 LQVLYPLSDPALNRLIHLVDRPAVLRSREQTAKALARDNGTVLGNTWYADPGLR 240

Qy 241 VVREVCRSVHRIAEBATLKPPFLGRYRLVRFEDLAREPLAEIRALYAFTGLSITPQLE 300
 Db 241 VVNEVCRSVHRIAEBALHKKPPFLQDRYRLVRFEDLARDPLTVIRELAFTGLGLTPQLO 300

Qy 301 AWIHNITHGSGPGRAREAFKTSRNALNVSQAWRHALPFKKIRRVQELCAGALQQLGYRP 360
 Db 301 TWINNITHGSGPGRAREAFKTSRNALNVSQAWRHTLPPAKIRRVQELCGALQQLGYRS 360

Qy 361 VYSEDEQRNIALDVLVRGLNGFTWASSTASHPRN 395
 Db 361 VHSLEQRDLSIDLULPRGMDSFKMASSTEKOPES 395

RESULT 12

Qy 301 AWIHNITHGSGPGRAREAFKTSRNALNVSQAWRHALPFKKIRRVQELCAGALQQLGYRP 360
 Db 301 TWINNITHGSGPGRAREAFKTSRNALNVSQAWRHTLPPAKIRRVQELCGALQQLGYRS 360

Qy 361 VYSEDEQRNIALDVLVRGLNGFTWASSTASHPRN 395
 Db 361 VHSLEQRDLSIDLULPRGMDSFKMASSTEKOPES 395

Query Match 74.8%; Score 1538; DB 16; Length 394;
 Best Local Similarity 93.3%; Pred. No. 7.2e-146; Indels 0; Gaps 0;
 Matches 294; Conservative 7; Mismatches 14;

Qy 1 MWLPRVSSTAVTALLAQTFLLFLVSRPGSSPAGGEARVHVLSSWRSGSSFVGOLF 60
 Db 1 MWLPRVSSTAVTALLAQTFLLFLVSRPGSSPAGGEARVHVLSSWRSGSSFVGOLF 60
 Qy 61 NQHPDVYLMEPAWHWTLLSQGSAATLHMAYRDLYRSVFLCDMDVFDAYLPQRNLSLD 120
 Db 61 NQHPDVYLMEPAWHWTLLSQGSAATLHMAYRDLYRSVFLCDMDVFDAYLPQRNLSLD 120

RESULT 13
 US-10-007-262-1
 ; Sequence 1, Application US/10007262
 ; Publication No. US2002016474A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bistrup, Annette
 ; APPLICANT: Rosen, Steven D.
 ; APPLICANT: Tangemann, Kirsen
 ; APPLICANT: Hemmerich, Stefan
 ; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
 ; FILE REFERENCE: 6510-107CIP
 ; CURRENT FILING DATE: 2001-11-08
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 386
 ; TYPE: PRT
 ; ORGANISM: H. sapiens
 US-10-007-262-1

Query Match 49.6%; Score 1019; DB 13; Length 386;
 Best Local Similarity 54.5%; Pred. No. 1.4e-91;
 Matches 216; Conservative 45; Mismatches 107; Indels 28; Gaps 7;

Qy 14 LLLAQTFLLFLVSRP-----GPSSPAGGEARVHVLSSWRSGSSFVGQL 59
 Db 2 LLPKMKMLLFLVSQLMALLPPHMYSNHSNISLMSMKAQPERMVLSSWRSGSSFVGQL 61
 Qy 60 ENQHQPDYFLMBPAWHTTLSQGSAATLHMAYRDLYRSVFLCDMDVFDAYLPQRNLS 118
 Db 62 FGQHPDYFLMBPAWHTTMSKOSTAAMLHMAYRDLYRAVFLCDMSYDMEPGRQS 121

Qy 14 LLLAQTFLLFLVSRP-----GPSSPAGGEARVHVLSSWRSGSSFVGQL 59
 Db 119 DLFOQWAWSRALCSPACAFPRGAISSEAVCKPLCARQSFTLAREACRSYSHVVLKEVR 178
 Qy 122 SLFQWENSRALCSAPACDIIPQDEIIIFRAHCRLLCSQQPEVEKACRSYSHVVLKEVR 181

Qy 179 FNQVQLYPLLSDFALNTRIVHLVRDPAVRLSREQTAKALARNDGTVLGNTGTW-BADP 237
 Db 182 FNQLQSLPLKQDSLNHIVHLVRDPAVRSRERTKGDLMSRIMGQHBOQLKEDQ 241

Qy 238 GLRVVREVCRSHVRIARAAATLKP-PFFLGRGRYLVRFEDLAREPLAIRALYAFTGLSLT 296
 Db 242 PYXQMVQTCOSOLEIYK-TIOSLPKALQERYLLVRYEDLARAPVATSRMVFVGLEFL 299

Query Match 49.6%; Score 1019; DB 9; Length 386;
 Best Local Similarity 54.5%; Pred. No. 1.4e-93; Indels 28; Gaps 7;
 Matches 216; Conservative 45; Mismatches 107;

Qy 14 LLLAQTFLLFLVSRP-----GPSSPAGGEARVHVLSSWRSGSSFVGQL 59
 Db 2 LLPKMKMLLFLVSQLMALLPPHMYSNHSNISLMSMKAQPERMVLSSWRSGSSFVGQL 61
 Qy 60 ENQHQPDYFLMBPAWHTTLSQGSAATLHMAYRDLYRSVFLCDMDVFDAYLPQRNLS 118
 Db 62 FGQHPDYFLMBPAWHTTMSKOSTAAMLHMAYRDLYRAVFLCDMSYDMEPGRQS 121

Qy 119 DLFOQWAWSRALCSPACAFPRGAISSEAVCKPLCARQSFTLAREACRSYSHVVLKEVR 178
 Db 122 SLFQWENSRALCSAPACDIIPQDEIIIFRAHCRLLCSQQPEVEKACRSYSHVVLKEVR 181

Qy 179 FNQVQLYPLLSDFALNTRIVHLVRDPAVRLSREQTAKALARNDGTVLGNTGTW-BADP 237
 Db 182 FNQLQSLPLKQDSLNHIVHLVRDPAVRSRERTKGDLMSRIMGQHBOQLKEDQ 241

Qy 238 GLRVVREVCRSHVRIARAAATLKP-PFFLGRGRYLVRFEDLAREPLAIRALYAFTGLSLT 296
 Db 358 GLRVVREVCRSHVRIARAAATLKP-PFFLGRGRYLVRFEDLAREPLAIRALYAFTGLSLT 296

RESULT 14
 US-10-841-707-6
 ; Sequence 6, Application US/10841707
 ; Publication No. US2004020264A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fukuda, Minoru

APPLICANT: Yeh, Jiunn-Chern
 APPLICANT: Hirao, Nobuyoshi
 TITLE OF INVENTION: Identification of the Meca-79 Antigen
 TITLE OF INVENTION: and Related Methods of Treating L-Selectin-Mediated
 Conditions
 FILE REFERENCE: P-LJ 4149
 CURRENT APPLICATION NUMBER: US/10/841,707
 CURRENT FILING DATE: 2004-05-06
 PRIOR APPLICATION NUMBER: US/09/569,320A
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 6
 LENGTH: 380
 TYPE: PRT
 ORGANISM: Homo sapien
 US-10-841-707-6

Query Match 49.5%; Score 1018; DB 16; Length 380;
 Best Local Similarity 55.0%; Pred. No. 1.7e-93; Gaps 7;
 Matches 214; Conservative 44; Mismatches 103; Indels 28; Gaps 7;

Qy 21 LLLFLYSLRP-----GPSSPAGEARVHVLVSSWRSGSSTVGOLFNOHDV 66
 Db 3 LLLFLYSLVSNVILALFFHMYSNINSSLSMKAPERMMHVLLVSSWRSGSSTVGOLFQOHDDV 62
 Qy 67 FYLMEPAWHTWTTLSQGSATLHNAVRDVLRSVFLCDMDVFDAYL-PWFRNLSDLFQWAV 125
 Db 63 FYLMEPAWHTWTTLFQOSTAMLNHNAVRDVLRAVFLCDMSVDAVMEPGPRQSSLFQWEN 122
 Qy 126 SRALCSPPACSAFPGAISSAEAVCKPLCARQSPTLAREACRSYSHVTLKEVRPFENLQVLY 185
 Db 123 SRALCSAPACDIPODEIIPRAHCRLLCSCQQPEVVERAKRSYSHVTLKEVRPNLQSY 182
 Qy 186 PLLSDPALNRIVHLVDPAVLRSRECTAKALARNDGVLTGNGTVW-BADPLRVIRE 244
 Db 183 PLLKDPDSLNLHIVHLVDPAVEFSRERTKGDLMDSRIVMGQHEQKLKKEDQPYVYNNQV 242
 Qy 245 VCRSHVRIAAATLKP-PPLFLRGYRLVRFEDLAREPLAERIALYAFTGLSLTQLEAWI 303
 Db 243 ICQSLQELEYIK-TIOSLPKALQERYLIVRLVREDLARAPVQTSRMVEFVGLFLPHLQTWV 300
 Qy 304 HNITHGSGPGRARFAFKTSSRNALNVSQWRHALPFAKTRVQELCAGAQOLIGYRPPYS 363
 Db 301 HNITRGKGGMG--DIAFHNTNARDLNQSRQLQACGDANLNGYRHS 358
 Qy 364 EDEQRNLALLDVLPRGLNGTWASSTASH 392
 Db 359 EDEQRNLALLDVLPRGLNGTWASSTASH 380

Search completed: June 23, 2005, 09:27:57
 Job time : 51.935 secs

RESULT 15

US-10-427-631-11
 Sequence 11, Application US/10447631
 Publication No. US20030175323A1
 GENERAL INFORMATION:
 APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
 APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
 APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
 APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
 APPLICANT: AZIMZAI, Valda
 TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 FILE REFERENCE: PF-0532-1 DIV
 CURRENT APPLICATION NUMBER: US/10/427,631
 CURRENT FILING DATE: 2003-04-29
 PRIOR APPLICATION NUMBER: US 09/786,240
 PRIOR FILING DATE: 2002-03-12
 PRIOR APPLICATION NUMBER: PCT/US99/20989
 PRIOR FILING DATE: 1999-09-09
 PRIOR APPLICATION NUMBER: US 60/172,220
 PRIOR APPLICATION NUMBER: US 60/155,248
 PRIOR FILING DATE: 1998-11-04

THIS PAGE BLANK (USPTO)

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--|
| 1 | 656.5 | 31.9 | 484 | 2 JB0261 | N-acetylglucosamin N-acetylglucosamin N-acetylglucosamin N-acetylglucosamin N-acetylglucosamin |
| 2 | 618.5 | 30.1 | 484 | 2 JC7350 | chondroitin 6-sulf chondroitin 6-sulf chondroitin 6-sulf chondroitin 6-sulf chondroitin 6-sulf |
| 3 | 615 | 29.9 | 486 | 2 A57351 | hypothetical prote probable enzyme, C probable ATP-binding protein |
| 4 | 532.5 | 25.9 | 458 | 2 A57397 | malate synthase (S hypothetical prote 3-deoxy-D-manno-octopulic acid esterase) |
| 5 | 128 | 6.2 | 398 | 2 G70729 | probable long-chain ferrous iron trans hypothethical prote |
| 6 | 118 | 5.7 | 307 | 2 B83394 | phospho-related prote |
| 7 | 100.5 | 4.9 | 523 | 2 B83629 | glucosamine-fructo mitochondrial inter |
| 8 | 98 | 4.8 | 559 | 2 T03412 | ATP-dependent Clp |
| 9 | 93 | 4.5 | 359 | 2 T16350 | hypothetical prote |
| 10 | 91.5 | 4.5 | 425 | 2 B83023 | polyketide synthase |
| 11 | 91.5 | 4.5 | 577 | 2 B75585 | acetolactate synthase |
| 12 | 91.5 | 4.5 | 617 | 2 F82744 | TRM3 transposase |
| 13 | 91.5 | 4.5 | 820 | 2 H86246 | probable transposase |
| 14 | 91 | 4.4 | 380 | 2 A75328 | TRM3 transposase |
| 15 | 90.5 | 4.4 | 349 | 2 E75617 | TRM3 transposase |
| 16 | 90.5 | 4.4 | 710 | 1 A462273 | TRM3 transposase |
| 17 | 90 | 4.4 | 931 | 2 AF32276 | TRM3 transposase |
| 18 | 89.5 | 4.4 | 221 | 2 T50665 | probable transposase |
| 19 | 89.5 | 4.4 | 668 | 2 T01685 | TRM3 transposase |
| 20 | 89.5 | 4.4 | 1489 | 2 S73015 | TRM3 transposase |
| 21 | 89.5 | 4.4 | 3643 | 2 T36410 | TRM3 transposase |
| 22 | 89 | 4.3 | 571 | 2 AG3174 | TRM3 transposase |
| 23 | 88.5 | 4.3 | 262 | 2 G95327 | probable transposase |
| 24 | 88.5 | 4.3 | 400 | 2 F42759 | probable transposase |
| 25 | 88.5 | 4.3 | 400 | 2 A42727 | probable transposase |
| 26 | 88.5 | 4.3 | 400 | 2 C95306 | probable transposase |
| 27 | 88.5 | 4.3 | 400 | 2 E95287 | probable transposase |
| 28 | 88.5 | 4.3 | 400 | 2 H95291 | probable transposase |
| 29 | 88.5 | 4.3 | 400 | 2 F95354 | probable transposase |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using BW model

Run on: June 23, 2005, 08:33:38 ; Search time 8.95931 Seconds (without alignments)
4242.027 Million cell updates/sec

Title: US-10-697-828-13
Perfect score: 2056
Sequence: 1 MWLPRVSSTAVTLLAQTF.....LPRGLNGFTWASSTASHPRN 395

Scoring table: BLOSUM62 Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

RESULT 1
JB0261
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
C.Species: Homo sapiens (man)
C.Date: 05-Feb-1999 #Sequence_revision 05-Feb-1999 #text_change 01-Mar-2002
C.Accession: JE0261
R.Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, Q.W.; Mitsuo, T.
J. Biochem. 124, 670-678, 1998
A.Title: Human N-Acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of
A.Reference number: JB0261; MUD:98:91845; PMID:9722682
A.Molecule type: mRNA
A.Residues: 1-484 <UCH>
A.Cross-references: DDBJ:AB014679
C.Comment: This protein catalyzes the transfer of sulfate from 3'-phosphoadenosine to the 6-OH group of N-acetylglucosamine.
C.Superfamily: chondroitin 6-sulfotransferase
C.Keywords: sulfotransferase

RESULT 1
JB0261
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
C.Species: Homo sapiens (man)
C.Date: 05-Feb-1999 #Sequence_revision 05-Feb-1999 #text_change 01-Mar-2002
C.Accession: JE0261
R.Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, Q.W.; Mitsuo, T.
J. Biochem. 124, 670-678, 1998
A.Reference number: JB0261; MUD:98:91845; PMID:9722682
A.Molecule type: mRNA
A.Residues: 1-484 <UCH>
A.Cross-references: DDBJ:AB014679
C.Comment: This protein catalyzes the transfer of sulfate from 3'-phosphoadenosine to the 6-OH group of N-acetylglucosamine.
C.Superfamily: chondroitin 6-sulfotransferase
C.Keywords: sulfotransferase

Query Match Score 656.5; DB 2; Length 484;
Best Local Similarity 39.0%; Pred. No. 5.9e-50;
Matches 146; Conservative 69; Mismatches 124; Indels 35; Gaps 10;

QY SPAG-GEARVHVHLVLSWRSGSSFVGCGLNQNPDVFLVMEPAFWHNTLSQGSGRAATHLIMA 91
Db 11.0 APGSVGDRHMMVYFTWRSGSFEGLFQNQNPBVFLPEWHRWQKLYPGDVAISLOGA 169

QY 92 VRDLVRSVFLCLMDVFDAYLP--WRMLNSLD--FQNAVSRALCSPACSAFPRGAIIS-- 144
Db 17.0 ARDMLSLAVALYRCDSVFQVLYSPAGSGGRLNLTTLGIFGAATNKVYCSSLCPAYRKEVVGLV 229

QY 145 SBAVCKPLCARQSFILAREACRYSHVYLKEVRFNLOVLYPLUSDPAINLRVHLVRDP 204
Db 23.0 DDRVCKK-CPPQLARFEECRKYRTVKGTRVEDAVLPLRDPALDKVHLVRDP 288

QY 205 RAVLRSROTAKALDRN-----GIVIGT--NCTWVEAD--PGIRV 241
Db 289 RAVASSRTRSRHLLIRESLQWVRSRDPBAHMPFLEAAGHKLGAKKEGVGGPDYHALCA 348

QY 242 VREVCRSHVRIAATLKPPFLRGRYKLVRFEDLARPLAEIRALVAFTGSLTTPQLEA 301
Db 34.9 MEVICNSWAKTLQTA-LQPPDMILQGHLYLVRVYDVLGVFPVKTLRRVYDFVGLLVSPENEQ 407

QY 302 WIRNITHESGPGRARFAFTSSRNALVNSQAWHAPLPAKIRVQEELCAGAOJLGVRPV 361
Db 408 FALMTSGSSSSK--PEVVSARNATQANANRATLFFQIKOVOEFCYQPMAVLGTERV 465

QY 362 YSEDEQRNIALDLY 375
Db 466 NSPEEVKOLSKTL 479

RESULT 2

JC7350 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
C:Accession: JC7350
R.Uchimura, K.; Fasakhany, F.; Kadomatsu, K.; Yamakawa, T.; Kurosoawa, N.; Blochim. Biophys. Res. Commun. 274, 291-296, 2000
A.Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
A.Reference number: JC7350; MUID:20374462; PMID:10913333
A.Molecule type: mRNA
A.Residues: 1-484 <UCH>
A.Cross-references: UNIPROT:Q93N00; UNIPROT:Q9EP78; DDBJ:AB040710
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylat sulfat linkage formation. This enzyme is involved in synthesis of L-selectin ligand in
C:Keywords: glycolysis; sulfat transferase; sulfotransferase; transmembrane protein

Query Match Score 618.5; DB 2; Length 484;
Best Local Similarity 38.0%; Pred. No. 1..3e-46;
Matches 148; Conservative 56; Mismatches 130; Indels 55; Gaps 11;

Qy 31 PSSP---AGEXA---RVHVLVSSRSRSGSSFVGQLFNOHQPDVEYLMEPAWHWNTLSQL 82
Db 82 PGSPGNLSAVGEAVTOEIQHIIYHATWRTGSSFLGELFNOHQPDVFVLYEPWHMNLQALYP 141

Qy 83 GSATLHMARDLVRSVPLCDMDVFDAY-----LPWERNLSD--LFWAVSRALCS 131
Db 142 GDAESLQALRDMLRSLRCDFSVLRVLYAQPGDGERAPDSANLTAMLFWRTNKVIC 201

Qy 132 PPACSAFP-----GAI\$AVCKPLCARQSF\$TAREACRSHVYLKEYRFENIQLQYLP 186
Db 202 PPLCPA\$PAPRADVGLVEDKA-CESTCPVSPVSLRAEBCRKPVWVIKDVLDFLGIVLP 260

Qy 187 LLSDPALA\$RIVHLVDRPVALRSREQTAKALARDNGIVLGTN-----GTWVEA 235
Db 261 LLRDGLNIVKVKVOLFRPVALNSRLKSROGLIRESIQVLRTRQGDHFHRVLLAHGVDA 320

Qy 236 DPG-----LRYVREVCRSHVRIA\$AATLKPPFLGCRYRLVR\$FEDLARE 279
Db 321 RPGQQARALPS\$PAPRADFFLTS\$ALEVICAEAWLRL-LFT\$GPAPL\$RRLYR\$EDLVWQ 379

Qy 280 PLAEIRALYTAFTG\$LTPQLEA\$W\$HNTGSGP\$GAR\$AFT\$K\$P\$T\$A\$N\$V\$O\$A\$W\$H\$A\$LP\$ 339
Db 380 PQQRLLRLLFSS\$URTLAALDAFAFMTRGSAYGADR-PHLSARDAREAVHWRERLSQL 438

Qy 340 AKIRRVOELCAGALQAGLGYRPYSEDEOR 368
Db 439 EQVRQVETACAPAMRLAY-PRSGDERD 466

RESULT 3

JC7351 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - human
C:Species: Homo sapiens (man)
C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004
C:Accession: JC7351
R.Uchimura, K.; Fasakhany, F.; Kadomatsu, K.; Yamakawa, T.; Kurosoawa, N.; Blochim. Biophys. Res. Commun. 274, 291-296, 2000
A.Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
A.Reference number: JC7350; MUID:20374462; PMID:10913333
A.Molecule type: mRNA
A.Residues: 1-486 <UCH>
A.Cross-references: UNIPROT:Q75667; DDBJ:AB040711
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylat sulfat linkage formation. This enzyme is involved in synthesis of L-selectin ligand in
C:Keywords: glycolysis; sulfat transferase; sulfotransferase; transmembrane protein

Query Match Score 615; DB 2; Length 486;
Best Local Similarity 38.0%; Pred. No. 2..7e-46;
Matches 144; Conservative 53; Mismatches 128; Indels 54; Gaps 8;

Qy 29 PG\$SSPAGG\$EA----RVHVLVSSRSRSGSSFVGQLFNOHQPDVEYLMEPAWHWNTLSQLGS 84
Db 86 PSNLSGA\$VGAVSREKQHIVH\$A\$WRTQH\$FL\$G\$E\$F\$N\$Q\$B\$D\$V\$F\$LY\$E\$P\$W\$H\$M\$A\$LY\$P\$D 145

Qy 85 AATLHMARDLVRSVPLCDMDVFDAY-----LPWERNLSD--LFWAVSRALCSPP 133
Db 146 AESLQALRDMLRSLRCDFSVLRVLYAPPGDPAA\$PDTANLTAA\$FWRTNKVIC\$PP 205

Qy 134 ACSAPPRG---AIS\$SEAVCKPLCARQSF\$TAREACRSHVYLKEYRFENIQLQYLP\$ 189
Db 206 LCPGAPRARAEGVLFV\$TACERS\$PPVIA\$RALEACRKPVWVIKDVL\$GIVLP\$LLR 265

Qy 190 DPALNLRIVHLVDRPVALRSREQTAKALARDNGIVLGTN-----GTW\$EAD\$P\$G 238
Db 266 DPGLNIVKVKVOLFRPVALNSRLKSROGLIRESIQVLRTRQGDHFHRVLLAHGV\$G\$P 325

Qy 239 -----LRYVREVCRSHVRIA\$AATLKPP\$FLGCRYRLVR\$FEDLARE 279
Db 326 GQSRLPA\$PAPRADFFLTS\$ALEVICAEAWLRLF\$ARGA-----PAWLR\$YLR\$RYEDLV\$RQ 381

Db 326 GQSRLPA\$PAPRADFFLTS\$ALEVICAEAWLRLF\$ARGA-----PAWLR\$YLR\$RYEDLV\$RQ 381

RESULT 4

A57397 Chondroitin 6-sulfotransferase (EC 2.8.2.17) precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #Text_change 09-Jul-2004
C:Accession: A57397
R.Fukuta, M.; Uchimura, K.; Nakashima, K.; Kato, M.; Kimata, K.; Shinomura, T.; Habuchi, J. Biol. Chem. 270, 1875-1880, 1995
A:Title: Molecular cloning and expression of chick chondrocyte chondroitin 6-sulfotransferase
A:Reference number: A57397; MUID:95355490; PMID:7629189
A:Accession: A57397
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-458 <FUK>
A:Cross-references: UNIPROT:Q92179; GB:D49915; NID:9971262; PID:BA08655.1; PID:9971263
C:Superfamily: chondroitin 6-sulfotransferase
C:Keywords: sulfotransferase

Query Match Score 532.5; DB 2; Length 458;
Best Local Similarity 34.6%; Pred. No. 4.9e-39; Mismatches 63; Indels 33; Gaps 10;

Qy 5 RV\$STAVTALLAQ\$PFL\$FLV\$RGP\$SPAGGEARVHV\$VLS\$MRG\$SS\$FGOLF\$NQHP 64
Db 80 RQDPA\$FSQURTRL\$RN\$TQLO\$A\$G\$G\$IA\$AP--EPRRHVLLMA\$T\$RGS\$FV\$G\$F\$N\$Q\$G 136

Qy 65 DV\$YLM\$P\$AVH--WTT\$Q\$G\$A\$AT\$L\$H\$M\$A\$V--RDLVRSVFL\$CD\$MD\$V\$F\$DAY--\$P\$W\$R\$NL 117

Db 137 NIFY\$P\$EPL\$H\$IE\$V\$T\$F\$P\$G\$A\$V\$G\$A\$V\$Y\$D\$L\$Q\$Q\$G\$F\$H\$G\$E\$P\$E\$H\$T 196

Qy 118 SD\$F\$W\$AV\$S\$R\$AL\$C\$S\$P\$A\$S\$F\$P\$R\$G\$A\$S\$S\$A\$E\$A\$V\$C\$K\$P\$L\$C\$A\$R\$Q\$S\$T\$A\$R\$E\$A\$C\$S\$Y\$H\$V\$K\$Y\$R 177

Db 197 AA\$F\$R\$G\$S\$S\$IL\$C\$E\$P\$V\$S\$L\$K\$V\$E\$Y\$K\$H\$C\$N\$R\$C\$P\$N\$U\$T\$A\$E\$C\$R\$K\$H\$N\$A\$V\$T\$R 256

Qy 178 F\$N\$Q\$Y\$P\$L\$S\$D\$P\$A\$N\$R\$V\$H\$Y\$D\$P\$A\$Y\$R\$S\$R\$Q\$T\$A\$K\$A\$R\$D\$G\$V\$G\$T\$N\$G\$T\$W\$--\$V 233

Db 257 I\$Q\$LE\$F\$Q\$P\$A\$E\$D\$P\$R\$D\$L\$R\$I\$Q\$Y\$V\$D\$P\$A\$Y\$V\$S\$M\$V\$A\$F\$-----G\$K\$Y\$E\$W\$K\$W\$A 306

Qy 234 EAD\$P\$G\$H\$R\$----V\$R\$E\$V\$C\$H\$V\$R\$IA\$E\$A\$T\$K\$P\$P\$F\$L\$G\$R\$Y\$R\$Y\$R\$F\$E\$D\$L\$A\$R\$P\$E\$L\$A\$T\$R\$A\$Y\$ 288

Db 307 E\$G\$A\$P\$1\$Q\$D\$E\$V\$Q\$R\$Q\$G\$N\$C\$E\$-I\$R\$U\$A\$E\$G\$L\$Q\$P\$M\$L\$R\$G\$Y\$M\$V\$Y\$E\$D\$V\$A\$R\$P\$K\$A\$L\$E\$M\$Y 365

Qy 289 AFT\$G\$SL\$T\$P\$Q\$LE\$A\$N\$T\$H\$T\$G\$G\$P\$G\$A\$R\$E\$A\$F\$K\$T\$S\$R\$N\$A\$N\$V\$Q\$A\$W\$H\$A\$L\$P\$F\$A\$K\$T\$R\$Q\$Q\$E\$L 348

| | | | | | |
|----------|--------|---|--|---|--|
| Db | 366 | RFAIHPHQVEEIRANT-.QAFQDSNGIYST-QKNSSEQFERWRFSPFKLAQVYQDA | 422 | A;Cross-references: UNIPROT:Q92VG4; GB:AL591985; PIDN:CAC49141.1; PMID:915140626; | |
| | | | | A;Experimental source: strain 1021, megaplasmid pSTM8B R.Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Fedderspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. | |
| Qy | 349 | CAGAQLLGXRYPSYEDGPNLADDLV | 375 | Science 293, 668-672, 2001 | |
| | | | | A;Authors: Kahn, D.; Kahan, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelauze, Lebault, P.; Vandemboul, M.; Vorholter, P.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K. A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti. | |
| Db | 423 | CEPAMRLFGYKLASSAQELTNRSI-SLL | 449 | A;Reference number: A966039; MUID:21368234; PMID:11474104 | |
| | | | | A;Contents: annotation C;Genetics: | |
| RESULT 5 | G70729 | hypothetical protein Rv2267c - Mycobacterium tuberculosis (strain H37RV) | | A;Gene: SMB21237 A;Genome: plasmid | |
| | | | | Query Match 5.7%; Score 118; DB 2; Length 307; Best Local Similarity 28.1%; Prd. No. 0.011; Indels 30; Gaps 9; Matches 50; Conservative | |
| Qy | 191 | PALMLRIVTHLVRDPRAV--- | LRSREQTAKALARNDNGIVLGTGNTWEADPGCLRVRREVCA | 246 | |
| | | | | 191 PGIDMRVTHLVRDPRGVWSLIGYERDAK-----SGLOKEIKP-KSVERPFL 195 | |
| Db | 149 | PGIDMRVTHLVRDPRGVWSLIGYERDAK-----SGLOKEIKP-KSVERPFL 195 | | Qy 149 R-SHVRVIAE---AATLKPPPLFGRYRIVRFEDIARPLAEITALYAFGLSLTPQELEM 302 | |
| | | | | Db 196 RWSMVNLAVEYLRSRKLSGEKVNMR---VRYEDFASDPVAVMQIGTPELDLS-QVGTS 249 | |
| Qy | 303 | IHNITHSGSGPG--ARRAEFKTSSRNANLVSQAVRHAPFAKTRVQEICAGAQLOLGY | 358 | Qy 303 IHNITHSGSGPG--ARRAEFKTSSRNANLVSQAVRHAPFAKTRVQEICAGAQLOLGY 358 | |
| | | | | Db 250 LEN-GEAMPGHQVAGNLRMNASIANLKDTRTRMPARQQVSFQLGGWMRLRRGY 306 | |
| RESULT 7 | B83629 | Query Match 6.2%; Score 128; DB 2; Length 388; | Best Local Similarity 29.9%; Pred. No. 0.002; Indels 106; Gaps 18; | probable ABC transporter PA0136 [imported] - <i>Pseudomonas aeruginosa</i> | |
| | | | | C;Species: <i>Pseudomonas aeruginosa</i> C;Date: 15-Sep-2000 #Sequence_revision 15-Sep-2000 #text_change 09-Jul-2004 | |
| Qy | 43 | VLVLSWSSFSVYQLF--NQH-----PDVFVLYMEPAWHWVTTLSQGSAATHL | 89 | C;Accession: B83629 C;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Brzadl, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, N.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 | |
| | | | | Nature 406, 959-964, 2000 | |
| Db | 83 | IFIVCHWRRTTLLHELVVYDRHTGPTGYCLAPHFFELTE----WFAPYYBFVLSKH | 137 | A;Cross-references: UNIPROT:Q916Z6; GB:AE004451; PIDN:99945958; PIDN:AA035242 | |
| | | | | A;Experimental source: strain PA01, an opportunistic pathogen C;Genetics: | |
| Qy | 90 | MAVRDLVRSYFLCDMDVFDAYPWWRNLSLDFOWAVSRAICSCPPACSAFRGATSEAVC | 149 | A;Reference number: A82950; MUID:10984043 | |
| | | | | A;Status: preliminary A;Molecule type: DNA A;Residues: 1-523 <STO> | |
| Db | 138 | RA-----MDNMD--LSLHHPQDEFVWCM-QGIPSPVLTAPPN----- | 173 | A;Cross-references: UNIPROT:Q916Z6; GB:AE004451; PIDN:99945958; PIDN:AA035242 | |
| | | | | A;Experimental source: strain PA01, an opportunistic pathogen C;Genetics: | |
| Qy | 150 | KPLCRAQSFTLAREACRSYSHVLUKEVRFPLNQVLY-----PLSDPMLNRI----- | 197 | A;Reference number: A82950; MUID:10984043 | |
| | | | | A;Accession: B83629 | |
| Db | 174 | RPPQYEYVLEQVAPRELB--IWKRTLPRFVQOYVYERRKTVLKNPHTSPFKIVLEV | 231 | Query Match 6.2%; Score 128; DB 2; Length 388; | |
| | | | | Best Local Similarity 29.9%; Pred. No. 0.002; Indels 106; Gaps 18; | |
| Qy | 198 | -----VHLRDPYAVLRSSEQTAKALARNDNGIVLGTGNTWEADPGL--RVREVCRSH | 249 | probable ATP-binding component of ABC transporter PA0136 [imported] - <i>Pseudomonas aeruginosa</i> | |
| | | | | C;Species: <i>Pseudomonas aeruginosa</i> C;Date: 15-Sep-2000 #Sequence_revision 15-Sep-2000 #text_change 09-Jul-2004 | |
| Db | 232 | FPOQAKFIHTRDPYVVPYPSITHLKAHYRKGHLQQPT----FDGLDDKVVSTYDVY | 284 | C;Accession: B83629 C;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Brzadl, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, N.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 | |
| | | | | Nature 406, 959-964, 2000 | |
| Qy | 250 | VRIAFAAATLKPPEFLGRVLRVRFEDLAREPLAEITALYAFGLS-----LTPOLEAWHN | 305 | A;Cross-references: UNIPROT:Q916Z6; GB:AE004451; PIDN:99945958; PIDN:AA035242 | |
| | | | | A;Experimental source: strain PA01, an opportunistic pathogen C;Genetics: | |
| Db | 285 | RKLDSGRELVDPT---RFYELRYDLGPEGQFLRRLQYHLGJGDPECYPLRQYL-- | 338 | Query Match 4.9%; Score 100.5; DB 2; Length 523; | |
| | | | | Best Local Similarity 26.8%; Prd. No. 0.8%; Indels 33; Gaps 6; | |
| Qy | 306 | ITHGSGPQAREAKTSS----RNALNYSQAW-----RHALPKIR | 343 | probable sulfotransferase protein [imported] - <i>Sinorhizobium meliloti</i> | |
| | | | | C;Species: <i>Sinorhizobium meliloti</i> C;Date: 24-Aug-2001 #Sequence_revision 24-Aug-2001 #text_change 09-Jul-2004 | |
| Db | 339 | -----ADHADYCTNSQVLTVEKRAI-YDEHGEIIDIYGYDRTPEPARLR | 383 | R;Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001 | |
| | | | | Nature 406, 959-964, 2000 | |
| RESULT 6 | E95334 | Probable enzyme, C-terminal domain similar to sulfotransferase protein [imported] - <i>Sinorhizobium meliloti</i> | | C;Species: <i>Sinorhizobium meliloti</i> C;Accession: E95334 | |
| | | | | R;Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001 | |
| Qy | 70 | MEPAWHWTTLSQGSATIH--MAVRDLVRSVPLCDMDVFDAYLPWWRNLSDLFQWAVS | 70 | A;Cross-references: UNIPROT:Q916Z6; GB:AE004451; PIDN:99945958; PIDN:AA035242 | |
| | | | | A;Experimental source: strain PA01, an opportunistic pathogen C;Genetics: | |
| Db | 385 | LALAEEITTRFAVYKAPGAGAPARSLSGGLQKPLI-----GREBILQAPRLVVA | 432 | A;Status: preliminary A;Molecule type: DNA A;Residues: 1-523 <STO> | |
| | | | | Nature 406, 959-964, 2000 | |
| Qy | 433 | AHPIFWGV----DVGAAJTHRALDGTATLUVSEDDELPL-----LSDRI----- | 478 | A;Cross-references: UNIPROT:Q916Z6; GB:AE004451; PIDN:99945958; PIDN:AA035242 | |
| | | | | A;Experimental source: strain PA01, an opportunistic pathogen C;Genetics: | |
| Db | 479 | AACLSGRGLCPAVATASASPKQV | | RESULT 8 | |
| | | | | T03412 maize synthase (EC 4.1.3.2) - maize | |

C;Species: Zea mays (maize)
C;Date: 24-Mar-1995 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03412
R;Paeck, N.C.
submitted to the EMBL Data Library, September 1994
A;Reference number: Z14939
A;Accession: T03412
A;Description: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-559 <PAA>
A;Cross-references: UNIPROT:P49081; EMBL:L35914; NID:9532624; PIDN:95326
A;Experimental source: strain TX5855; scutellum
C;Genetics:
C;Note: MS
C;Function: <ACG>
A;Description: catalyzes the aldol condensation of glyoxylate with acetyl-CoA to form malate
C;Supfamily: glyoxylate cycle
C;Keywords: malate synthase
C;Keywords: carbon-carbon lyase; coenzyme A; oxo-acid-lyase

Query Match 4.8%; Score 98; DB 2; Length 559;
Best Local Similarity 24.7%; Pred. No. 1..5; Mismatches 54; Conservative 29; Indels 56; Gaps 12;
Matches 54;

Qy 160 LARBACTSYSHVVLKEVRFNLQVLLPPALNRIVLVRDPAVRSREQ 213
Db 313 MAQHEMRSYSHVLLHTCHHRGVHAGMGAAPKIDDAANAEALEVVKQL---REV 368
Qy 214 TAKALARNDGIVLGTGTWADPGL-RVREVCRSHV---RIAAATLKPPIPFLRGR 267
Db 369 RA-----GHDGTA-AHGLLPAIREVEFHIGGRPNQIGDAAGHEGAS----
Qy 268 YRLVRFEDLAREPLAEIRAYATGLSLTPQ---LEAWHNITHSGPGRAREEFKS 322
Db 412 ---VKEEDIQPP---RGARTYDGLRNLYVQYTLAAWL---AGGSYVPLYNLMEDA 460
Qy 323 SRLANVSOAW---RHALPF---AKTERVQELCAGALQ 354
Db 461 ATABISRYQNQWQMRHGAAALDAGSVEFATPEELARVVE 499

RESULT 9
T16350 hypothetical protein F42G9.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
R.Taich, A.
submitted to the EMBL Data Library, March 1996
A;Description: the sequence of C. elegans cosmid F42G9.
A;Reference number: Z18498
A;Accession: T16350
A;Molecule type: DNA
A;Residues: 1-359 <TAI>
A;Cross-references: UNIPROT:Q20351; EMBL:U00051; NID:91216305; PIDN:AAA913
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:F42G9.8
A;Introns: 28/2; 50/2; 113/2; 155/3; 220/3; 290/2
C;Supfamily: Caenorhabditis elegans hypothetical protein F42G9.8

Query Match 4.5%; Score 93; DB 2; Length 359;
Best Local Similarity 19.7%; Pred. No. 2..3; Mismatches 50; Conservative 50; Indels 106; Gaps 17;
Matches 66;

Qy 50 RSGSSFVGOLENOHPDV---FVLMEDAWHWWTLSQGSAAATHMAVRDLRSVFLCDMD 105
Db 95 RSGCTLDMPAIDLDAHPDVCGGETMLLSFLTVQA-----128
Qy 106 VFDAYLPWRRNLSDLFQAVASRALICS---PACCSAFPRGAISSEAVCKP-LCARQSPTLA 161
Db 129 -----GHRND-----WNNNSGITQEVFDAVSAFITEVAKHSSELAPRCLNKDPYTA- 175

Query Match 4.5%; Score 91.5; DB 2; Length 425;
Best Local Similarity 22.2%; Pred. No. 3..8; Mismatches 41; Conservative 41; Indels 127; Gaps 20;
Matches 88;

Qy 31 PSSPAGEGEARYHVVLSSNRSGSSPVGQFLPNOHPDV---FYLMEPAWYWTTLSQGAAAT 87
Db 44 PEVPPPGC-IWHAWSVGGESTAAPNPVALLERHPOLPVTCMTT-----GS--- 90

Qy 88 LHMAVRLDVLRSVFLCDMDVFDAYLPWPRNLSLDFOMAWSRALCSPACSAFPREGAISEA 147
Db 91 -----BRIRALF---GEQVRHCYLV-----DL-PWAARFL---DRYPRPLAVMET 131
Qy 148 VCKP---LCAROSF-----TLLAREACRSYHV---VLKEVRFNLUQ-----182
Db 132 ELWPNHIIHACAVRGIPVVALANARLERSARYARFGLTRPMLAEWSIAVQTEAEARF 191

Qy 183 -----VLYPLISDPALNLRIVLVRDPAVRSREQTAKALARNDGIVLG 227
Db 192 RRLGARPECVSVTGSISKEDRIDPOLPAAALRBWDATARPLMIASTHAGEDEIVLA 251

Qy 228 TNGTWBADDPLCRLVREVCRSHVIAEATLKP- PPFLRGRYRLVRFEDLAR---E 279
Db 252 AHRRLLETRPDALLI-----LVPRIHPERFQGHVHCRREGFATYRSGG 296

Qy 280 PLA-----BIALY-----AFTGSLITPQLEAMINHTHSGIG---AREA 318
Db 297 PVARATQVLGDTMGELFLYALADIAFVGSSLVN---GGHNLLEPAALGKPVFAGPHL 353

Qy 319 FKTSSRNALNVSQWHLPAFKTIRVOELCAGALQ 355
Db 354 P-----NFLDIAQLRDAGALLEVIDGELCDGLARL 385

chado, M.A.; Madeira, A.M.B.N.; Marinho, C.L.; Marques, M.V.; Martins, E.; Martins, B.M.; Martins, E.M.; Matsukuma, A.Y.; Menck, C.P.M.; Miracca, B.C.; Miyaki, C.Y.; Oliveira, M.C.; Oliveira, M.C.; Palmieri, D.A.; P.F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; Rosa Jr., V.E.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Silveira, S.; da Silveira, S.; da Silveira, S.; da Silva, A.C.R.; da Silva, F.R.; da Silva, J.R.; W.A.; da Silva, M.H.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z.; A; Reference number: A59328
A; Comments: annotation
C; Genetics:
A; Gene: Xf9933
C; Superfamily: ferrous iron transport protein B; translation elongation factor Tu homolog

| | | | | |
|----|-----|--|-----|--|
| Db | 263 | FYMDYRGSVIRR-----WSETRRNWTVGLOVPAATECDMYRRCGEFTCNPRKNPL | 313 | A;Title: Genome sequence of the radioresistant bacterium <i>Deinococcus radiodurans</i> R1. |
| Qy | 135 | C\$AF----PRAI-----SSAEACK---PL-CAROSFTLAREACRSYSHVILKEVRFNL | 181 | A;Reference number: A75250; MUID:20036898; PMID:10567266 |
| Db | 314 | C\$CIRGFRRPNLLEWNNGNWSGGCTTRVPLQCRQNNGS----- | 353 | A;Accession: E75611 |
| Qy | 182 | QVLYPLLSDPALNLRIVL-----VRDPRAVLRSREQTAKALARNDNGIVLGT--NG | 230 | A;Status: preliminary |
| Db | 354 | :-----ADGELFLRLRMKLPDFARRSESPEC-LRTLCQTCSCIAAHGLGCGMTIWNG | 405 | A;Molecule type: DNA |
| Qy | 231 | TWEADPGFLRVREVCRS---EVRIA-----EAATLKPPFLRG | 266 | A;Residues: 1-349 <WHI> |
| Db | 406 | SLVDS-----QELSAASGLDLXIRLANSEITKDKRPLIG | 440 | A;Cross-references: UNIPROT:Q9RYZ9; GB:AE001862; PIDN:AAF1221C |
| | | | | A;Experimental source: strain R1 |
| | | | | C;Genetics: |
| | | | | A;Gene: DR0154 |
| | | | | A;Map position: 2 |
| | | | | Query Match 4.4%; Score 90.5; DB 2; Length 349; |
| | | | | Best Local Similarity 24.1%; Pred. No. 3.6%; Mismatches 133; Indels 79; Gaps 15; |
| | | | | Matches 78; Conservative 34; MisMatch 34; |
| Qy | 33 | SPAGGERARVHLVLSSEPGSSSEVGOLFQNQHEDVYLM-----EAW-----H | 75 | Y;Score 90.5; DB 2; Length 349; |
| Db | 2 | TDASAPAPAEPLMLRENEAFAVVRLLNEQDAVARLANAIRERQAYAVTIARGSSDH | 61 | Best Local Similarity 24.1%; Pred. No. 3.6%; Mismatches 133; Indels 79; Gaps 15; |
| Qy | 76 | WTTLSQSQAATLHMAYRDLYRSVFLCDMDYFDAYLFWRNRNLSDFLFWAVSRALCSPPAC | 135 | Y;Score 90.5; DB 2; Length 349; |
| Db | 62 | ACTVLKALETOLSLPVASLGPSVH-----TLYGA-----RLDIAGALVAVSGASDVV | 113 | Best Local Similarity 24.1%; Pred. No. 3.6%; Mismatches 133; Indels 79; Gaps 15; |
| Qy | 136 | S---APRGAISSEAV--CKPLCARQSFTL-----AREACSYSHVILKEVRFEN | 180 | Y;Score 90.5; DB 2; Length 349; |
| Db | 114 | ENVMRAREGGALTVLVNDEPLAEEFTPLRCGPEKAVATKSY---LASLCAF- | 168 | Best Local Similarity 24.1%; Pred. No. 3.6%; Mismatches 133; Indels 79; Gaps 15; |
| Qy | 181 | LQVLYPLLSDPALNLRVHLDPRAVLRSREQTAKALA-----RDGIVLVLTGNTVTEA | 235 | Y;Score 90.5; DB 2; Length 349; |
| Db | 169 | LPLVLAELTGDEALK-RALDALPEQLAHTLTLEDAAHLLAELGARYRFADNLILARGHHYGV | 227 | Best Local Similarity 24.1%; Pred. No. 3.6%; Mismatches 133; Indels 79; Gaps 15; |
| Qy | 236 | DPLGRVYREVCRSHVIAEAATLKPPFLGRYRLV-----RFE | 274 | Y;Score 90.5; DB 2; Length 349; |
| Db | 228 | QEAAALKRKEETCGTHAEYSA-----FSHGPKRLLAEGPLPLGASAADAWEATRDAYD | 282 | Best Local Similarity 24.1%; Pred. No. 3.6%; Mismatches 133; Indels 79; Gaps 15; |
| Qy | 275 | DLAREPLAEIRALYAFFGLSL-TP | 297 | Y;Score 90.5; DB 2; Length 349; |
| Db | 283 | DL-RAAGDLRLIGPVGADLPTP | 305 | Best Local Similarity 24.1%; Pred. No. 3.6%; Mismatches 133; Indels 79; Gaps 15; |
| | | | | Search completed: June 23, 2005, 08:50:17 |
| | | | | Job time : 10.9593 secs |
| Qy | 124 | AVSPLCSPACSAFPRAISSEAVCKLICAROSFTLARE---ACRSYSHVILKEVRF | 179 | RESULT 14 |
| Db | 2 | ALSECFLSFP---FPPLRE-----CAPAGARTGDKLTDSQTNIAPPALDQTVTATLN | 55 | Y;Score 90.5; DB 2; Length 349; |
| Qy | 180 | NQVLYPLLSDPALNLRVHLDPRAVLRSREQTAKALARNDNGIVLTGNTWVE-ADP | 237 | Best Local Similarity 26.5%; Pred. No. 3.7%; Mismatches 104; Indels 60; Gaps 14; |
| Db | 56 | DOREAYALIGANDNLR-----RMRREULTRAKLARGEVTITGDAADVEGAER | 103 | Y;Score 90.5; DB 2; Length 349; |
| Qy | 238 | GLRVVREVCRSHVIAEAATLKPPFLGRYRLVRFEDAREPLAEIRALYAFFGLSL-- | 295 | Best Local Similarity 26.5%; Pred. No. 3.7%; Mismatches 104; Indels 60; Gaps 14; |
| Db | 104 | MVRDALDVTRS-----GGELTPDSLLS---ARLSSBGRSLAAETQV---NGLTLPR | 149 | Y;Score 90.5; DB 2; Length 349; |
| Qy | 296 | -----TPOLEAWI-----HNITHGSGGP GARAEFKTSSRNALNSQAWRHALPFKAIRR | 345 | Best Local Similarity 26.5%; Pred. No. 3.7%; Mismatches 104; Indels 60; Gaps 14; |
| Db | 150 | GLKEPKTGOKYLDDINESDTFEGPAG---TGKTYMAVAMAV-----QAJKAKYKRI | 201 | Y;Score 90.5; DB 2; Length 349; |
| Qy | 346 | --QELCAGAQOLQGYRP | 360 | Best Local Similarity 26.5%; Pred. No. 3.7%; Mismatches 104; Indels 60; Gaps 14; |
| Db | 202 | ILTRPAVENGGKLGFLP | 218 | Y;Score 90.5; DB 2; Length 349; |
| | | | | RESULT 15 |
| | | | | E75611 |
| | | | | glucosamine-fructose-6-phosphate aminotransferase-related protein - <i>Deinococcus radiodurans</i> |
| | | | | C;Species: <i>Deinococcus radiodurans</i> |
| | | | | C;Accession: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004 |
| | | | | C;Accession: E75611 |
| | | | | R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; |
| | | | | S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma |
| | | | | Science 286, 1571-1577, 1999 |

| Result No. | Score | Query | Match | Length | DB | ID | Description |
|------------|--------|-------|-------|--------|------------|----|----------------------------------|
| 1 | 205.6 | 100.0 | 395 | 2 | Q9GZK3 | | Q9gxx3 homo sapien |
| 2 | 1729.5 | 84.1 | 390 | 2 | Q7LCN3 | | Q7Lcn3 homo sapien |
| 3 | 1729.5 | 84.1 | 411 | 2 | Q9GZS9 | | Q9gzs9 homo sapien |
| 4 | 1708 | 83.1 | 395 | 2 | Q9QUP4 | | Q9qup4 mus musculus |
| 5 | 1024 | 57.2 | 392 | 2 | Q6GN39 | | Q6gn39 xenopus laevis |
| 6 | 1024 | 49.8 | 386 | 2 | Q8NCG5 | | Q8ncg5 homo sapien |
| 7 | 1019 | 49.6 | 386 | 2 | Q9Y5R3 | | Q9y5r3 homo sapien |
| 8 | 1012 | 49.2 | 370 | 2 | Q8IV46 | | Q8iv46 homo sapien |
| 9 | 984 | 47.9 | 388 | 2 | Q9RI11 | | Q9ri11 homo sapien |
| 10 | 977 | 47.5 | 388 | 2 | Q9WUE5 | | Q9wue5 mus musculus |
| 11 | 663 | 32.2 | 483 | 2 | Q94G9 | | Q94g9 mus musculus |
| 12 | 663 | 32.2 | 530 | 2 | Q8B276 | | Q8b276 mus musculus |
| 13 | 660 | 32.1 | 530 | 2 | Q80WV3 | | Q80wv3 mus musculus |
| 14 | 657.5 | 32.0 | 483 | 2 | Q9UED5 | | Q9ued5 homo sapien |
| 15 | 657.5 | 32.0 | 530 | 2 | Q9Y4C5 | | Q9y4c5 homo sapien |
| 16 | 622.5 | 30.3 | 484 | 2 | Q9EP78 | | Q9ep78 m-n-acetylglucosaminidase |
| 17 | 619.5 | 30.1 | 484 | 2 | Q9NB0 | | Q9nb0 mus musculus |
| 18 | 615 | 29.9 | 486 | 2 | Q75667 | | Q75667 homo sapien |
| 19 | 614 | 29.9 | 486 | 2 | Q9NS84 | | Q9ns84 homo sapien |
| 20 | 593 | 28.8 | 420 | 2 | Q6DBY9 | | Q6dby9 brachydanio |
| 21 | 593 | 28.8 | 485 | 2 | Q6XOG8 | | Q6xog8 rattus norvegicus |
| 22 | 578.5 | 28.1 | 479 | 2 | Q7LGCB | | Q7lgcb rattus norvegicus |
| 23 | 576.5 | 28.0 | 479 | 2 | Q7P099 | | Q7p099 homo sapien |
| 24 | 569 | 27.7 | 472 | 2 | Q88199 | | Q88199 mus musculus |
| 25 | 564.5 | 27.5 | 411 | 2 | Q9EQCO | | Q9eqco mus musculus |
| 26 | 555 | 27.0 | 411 | 2 | Q79415 | | Q79415 homo sapien |
| 27 | 555 | 27.0 | 474 | 2 | Q9QZL2 | | Q9qzl2 rattus norvegicus |
| 28 | 532.5 | 25.9 | 458 | 1 | C6ST CHICK | | C6st chick |
| 29 | 463.5 | 22.5 | 441 | 2 | Q79403 | | Q79403 torpedo catfish |
| 30 | 393.5 | 19.1 | 257 | 2 | Q79415 | | Q79415 mus musculus |
| 31 | 347.5 | 16.9 | 225 | 2 | Q6RY62 | | Q6ry62 cavia porcellus |

| Scoring table: BLASTM62 Gapop 10.0 , Gapext 0.5 | | | | | | |
|---|--|--|--|--|--|--|
| Searched: 1612378 seqs, 512079187 residues | | | | | | |
| Total number of hits satisfying chosen parameters: 1612378 | | | | | | |
| Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries | | | | | | |
| Database : UniProt 03: 1: uniprot_sprot;* 2: uniprot_trembl;* | | | | | | |
| Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. | | | | | | |
| SUMMARIES | | | | | | |
| RESULT 1 Q9GZK3 ID Q9GZK3 ; PRELIMINARY; AC Q9GZK3 ; Created) DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update) DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update) DE N-acetylglucosamine 6-O-sulfotransferase GST-4beta (Corneal N-acetylglucosamine-6-O-sulfotransferase) (Carbohydrate (N-Glycan) beta 1,6-N-acetylgalactosaminidase) DE Name=GST4beta; Synonyms=CHST6; GN Homo sapiens (Human) OS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Euarchontoglires; Aves; Cetartiodactyla; Primates; Catarrhini; Hominidae; Homo sapiens OC NCBI_TAXID=9606; RN [1] RP SEQUENCE FROM N.A. RX MEDLINE=21056027; PubMed=11181564; DOI=10.1093/glycob/cuq075; RA Hammerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddie N.R., Rosen S.D.; RA RT "Chromosomal localization and genomic organization of the galactose/N-acetylglucosamine/N-acetylglucosamine 6-O-sulfotransferase gene"; RT Glycobiology 11:75-87(2001). RN [2] RP SEQUENCE FROM N.A. RX MEDLINE=20472330; PubMed=11017006; DOI=10.1038/79987; RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Ozaki K., Nakamura T., Dots A., Kawasaki S., Inoue Y., Maeda N., Yamamoto S., Fujiwara T., Thomar E.J., Shimomura Y., Kinoshita S., Tanigami A., Fukuda M.N.; RA RT "Macular corneal dystrophy type I and type II are caused by distinct mutations in a new sulphatotransferase gene."; RT Nat. Genet. 26:237-241(2000). RN [3] RP SEQUENCE FROM N.A. RX TISSUE=Ductal; RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Schueler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaeffer C.M., Schaeffer C.M., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marubina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C., Raha S., Loqueland N.A., Peters G.J., Abramson R.D., Mullany S.J., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villaillon D.K., Muñoz D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Heitton E., Kettman M., Madan A., Rodriguez S., Sanchez A., RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Schatz J., Myers R.M., Butterfield Y.S., Rodriguez A.C., Grimwood J., Schatz J.E., Krzywinski M.I., Skalska U., Smailus D.E., Schein J.E., Schein J.E., RA | | | | | | |

| | | | | |
|-----------------------|--|--|--|---|
| RA | Jones S.J., Marra M.A., | Analysis of more than 15,000 full-length human and mouse cDNA sequences.", | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). | [4] |
| RN | RP | SEQUENCE FROM N.A. | | |
| RC | TISSUE=Lung; | | | |
| RA | Director MGC Project; | | | |
| RL | Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; AF280086; AAG48244.1; -. | | | |
| DR | EMBL; AF26127.1; -. | | | |
| DR | EMBL; BC074883; AAH74883.1; -. | | | |
| DR | EMBL; BC074834; AAH74834.1; -. | | | |
| DR | EMBL; AF219980; AAG26125.1; -. | | | |
| DR | GeneID: HGNC:6938; CHST6. | | | |
| DR | GO: GO:005794; C: Golgi apparatus; TAS. | | | |
| DR | GO: GO:0006044; P:N-acetylglucosamine metabolism; IC. | | | |
| DR | InterPro; IPR00863; Sulfotransferase. | | | |
| DR | Pfam; PF0685; Sulfotransferase_1. | | | |
| KW | Transfase. | | | |
| SQ | SEQUENCE 395 AA; 44098 MW; 433CA60249A48F67 CRC64; | | | |
| Query Match | 100.0%; | Score 2056; DB 2; Length 395; | | |
| Best Local Similarity | 100.0%; | Pred. No. 1.e-165; | | |
| Matches | 395; | Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | |
| Qy | 1 MWLPRVSTAVTALLQ-TFLLIPLVSREGPSSPAGEARVHLVLSWRGSSFVGQLF | 60 | Qy | 1 MWLPRVSTAVTALLQ-TFLLIPLVSREGPSSPAGEARVHLVLSWRGSSFVGQL 59 |
| Db | 1 MWLPRVSTAVTALLQ-TFLLIPLVSREGPSSPAGEARVHLVLSWRGSSFVGQLF | 60 | Db | 1 MWLPRVSTAVTALLQ-TFLLIPLVSREGPSSPAGEARVHLVLSWRGSSFVGQL 60 |
| Qy | 61 NQHQDVFYLMPEAHWVNTLSSQGSAATLHMARDLVRSVFLCDMDVFDAYLPWRNLSDL | 120 | Qy | 60 FNOHQDVFYLMPEAHWVNTLSSQGSAATLHMARDLVRSVFLCDMDVFDAYLPWRNLSDL 119 |
| Db | 61 NQHQDVFYLMPEAHWVNTLSSQGSAATLHMARDLVRSVFLCDMDVFDAYLPWRNLSDL | 120 | Db | 61 FSQHQDVFYLMPEAHWVNTLSSQGSAATLHMARDLMSFLCDMDVFDAYMPQRNLSDA 120 |
| Qy | 121 FQFAVSRALCSPPACSAFPGRATSSAEVAKPLCARQSFTLAREACRSTSHTVKEVRFEN | 180 | Qy | 120 LFQWAVSRALCSPPACSAFPGRATSSAEVAKPLCARQSFTLAREACRSTSHTVKEVRF 179 |
| Db | 121 FQFAVSRALCSPPACSAFPGRATSSAEVAKPLCARQSFTLAREACRSTSHTVKEVRFN | 180 | Db | 121 FFNWAVSRALCSPPACSAFPGRATSSAEVAKPLCARQSFTLAREACRSTSHTVKEVRF 180 |
| Qy | 181 LQVLYPLISDPALNLRVHLVDPRAVLRSQTAKALRDVRSVFLCDMDVFDAYLPWRNLSDL | 240 | Qy | 180 NLQVLYPLISDPALNLRVHLVDPRAVLRSREQTAKALRDVRSVFLCDMDVFDAYLPWRNLSDL |
| Db | 181 LQVLYPLISDPALNLRVHLVDPRAVLRSQTAKALRDVRSVFLCDMDVFDAYLPWRNLSDL | 240 | Db | 181 NLQVLYPLISDPALNLRVHLVDPRAVLRSREAGPILLARDVNGIVLGTGKWTADPHL 240 |
| Qy | 241 WREVCRSHVRIBAAATLKPPLPGLYRVLRFEDLAPEIRALYAFGLSLPQLE | 300 | Qy | 240 RVREVCRSHVRIBAAATLKPPLPGLYRVLRFEDLAPEIRALYAFGLSLPQLE |
| Db | 241 WREVCRSHVRIBAAATLKPPLPGLYRVLRFEDLAPEIRALYAFGLSLPQLE | 300 | Db | 241 RLIRECRSHVRIBAAATLKPPLPGLYRVLRFEDLAPEIRALYAFGLSLPQLE |
| Qy | 301 AWIHNITHSGGP GARREAFKTSSRNALNVSOAWRHALPFAKIRVOELCAGALQLGYR | 360 | Qy | 300 EAWIHNITHSGGP GARREAFKTSSRNALNVSOAWRHALPFAKIRVOELCAGALQLGYR |
| Db | 301 AWIHNITHSGGP GARREAFKTSSRNALNVSOAWRHALPFAKIRVOELCAGALQLGYR | 360 | Db | 301 EAWIHNITHSGGP GARREAFKTSSRNALNVSOAWRHALPFAKIRVOELCAGALQLGYR |
| Qy | 361 VVSEDEQPNLADLVLRGLNGFTWASSTASHPRN 395 | 360 | Qy | 360 PVYSEDEQPNLADLVLRGLNGFTWAS 387 |
| Db | 361 VVSEDEQPNLADLVLRGLNGFTWASSTASHPRN 395 | 360 | Db | 361 PVYSEADQQRDLTDLVLRGPDHFSWAS 388 |
| RESULT 3 | | | | |
| Q7LCN3 | PRELIMINARY; | PRT; 411 AA. | Q9GZS9 | PRELIMINARY; |
| ID | Q7LCN3; | | Q9GZS9; Q9UBY3; | |
| AC | | | AC Q9GZS9; Q9UBY3; | |
| DT | 05-JUL-2004 (TREMBrel. 27, Created) | | DT 01-MAR-2001 (TREMBrel. 16, Created) | |
| DT | 05-JUL-2004 (TREMBrel. 27, Last sequence update) | | DT 01-MAR-2001 (TREMBrel. 16, Last sequence update) | |
| DT | 05-JUL-2004 (TREMBrel. 27, Last annotation update) | | DT 05-JUL-2004 (TREMBrel. 27, Last annotation update) | |
| DE | N-acetylglucosamine 6-O-sulfotransferase. | | DE 6-O-sulfotransferase (Intestinal N-acetylglucosamine- | |
| DE | Name=1-GlcNAc-6-ST; | | DE Name=CHS5; | |
| OS | Homo sapiens (Human). | | OS Homo sapiens (Human). | |
| OS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OC | | | GN Homo sapiens (Human). | |
| OC | | | NCBI TaxID=9606; | |
| OX | NCBI TaxID=6006; | | RN Akama T.O., Nishida K., Nakayama J., Watanabe H., Ozaki K., Nakamura T., Dots A., Kawasaki S., Inoue Y., Maeda N., Yamamoto S., Fujiwara T., Thonar E.J., Shimomura Y., Kinoshita S., Tanigami A., Fukuda M.N., | |
| RN | | | RA "Macular corneal dystrophy type I and type II are caused by distinct | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Intestine | | | |
| RX | MEDLINE=99423499; PubMed=10491328; DOI=10.1006/bbrc.1999.1324; | | | |

| | |
|-----------------------|--|
| RT | mutations in a new sulphotransferase gene."; |
| RL | Genet. 26:237-241(2000). |
| DR | EMBL; AF246718; AAG28033.1; -. |
| DR | EMBL; AF219991; AAG2626.1; -. |
| DR | Gene; HGNC:1973; CHST5. |
| GO | GO:0008146; F:sulphotransferase activity; IEA. |
| GO | GO:0016740; F:transferase activity; IEA. |
| DR | InterPro; IPR000563; Sulfortransferase. |
| DR | PF00685; Sulfortransferase_1. |
| KW | Sulphotransferase; Sulfortransferase. |
| SEQUENCE | 411 AA; 46160 MW; 97642D54BE926E06 CRC64; |
| Query Match | Score 1729.5; DB 2; Length 411; |
| Best Local Similarity | 85.8%; Pred. No. 7.e-137; |
| Matches | 333; Conservative 17; Mismatches 37; Indels 1; Gaps 1; |
| Qy | 1 MWLPRVSSTAVTALLAQ-TFLILFLVSRGPSSPAGGEARHFLVLSSWRGSSFWQL 59 |
| Db | 22 MWLPRFSSKTTVTILLAQTCLLFLISRGPSPPAGEDRVRVLSSWRGSSFWQL 81 |
| Qy | 60 FNQHPDVYLMPEPAWHVWTLSQGSAAATHMAYRDLMRFDAYLPWRRNLSD 119 |
| Db | 82 FSQHPDVYLMPEPAWHVWTLSQGSAAATHMAYRDLMRFDAYLPWRRNLSD 141 |
| Qy | 120 LFQWAVSRLCSPPACSAFPGRGAISSCAEVCKPLCARQSTFLAREACRSYSHVILKEVRFN 179 |
| Qy | 142 FFWATSLCSPPACSAFPGRGTSKQDVCTKLTQFSLAREACRSYSHVILKEVRF 201 |
| Db | 180 NQVLYPLISDPALNLRVHLVEDPRAVLSREOTAKALARNDIVLGNTGWAEADPLR 239 |
| Qy | 202 NLQVLYPLISDPALNLRVHLVRDPRAVLSREAGPILARNDIVLGNTGNRVEADPHL 261 |
| Qy | 240 RYREVCRSHVRLAEEATLKPPPLGRYRVRFEDLAREPLAEIRALYAFTGLSLTPQL 299 |
| Db | 262 RLREVCRSHVRLAEEATLKPPPLGRYRVRFEDLAREPLAEIRALYAFTGLSLTPQL 321 |
| Qy | 300 EAHTNITHSGSGPGRARRAKFTSSRNALNSQMRHALPKAKRVOELCAGALQLGYR 359 |
| Db | 322 EAHTNITHSGGKPKIAFTHFTSSRNALNSQMRHALPKAKRVOELCAGALQLGYR 381 |
| Qy | 360 PVSDEORNLADLVLPGLNGTWAS 387 |
| Db | 382 PVYSADQQRDLTDLVLPQGDHFWSAS 409 |
| RESULT 4 | |
| ID | Q9QUP4 |
| AC | Q9QUP4, |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) |
| DT | 01-MAY-2000 (TREMBLrel. 13, Last sequence update) |
| DT | 05-JUL-2004 (TREMBLrel. 27, Last annotation update) |
| DE | "Cloning and characterization of a mammalian N-acetylglucosamine-6-sulphotransferase that is highly restricted to intestinal tissue." |
| GN | Name=Chat5; Synonyms=I-GlcNAc-6-ST; |
| OS | Mus musculus (Mouse). |
| OC | Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Murae; Murinae; Mus. |
| RN | NCBI_TaxID=10090; |
| OX | NCBI_TaxID=10090; |
| RP | SEQUENCE FROM N.A. |
| RC | SPAIN-E57B1/6; TISSUE=Intestine; MEDLINE=91942499; PubMed=1091328; DOI=10.1006/bbrc.1999.1324; |
| RX | Lee J.K., Bhakta S., Rosen S.D., Hemmerich S.; "Cloning and characterization of a mammalian N-acetylglucosamine-6-sulphotransferase that is highly restricted to intestinal tissue." Biochem. Biophys. Res. Commun. 263:543-549 (1999). |
| RA | [2] |
| RN | SEQUENCE FROM N.A. |
| RC | SPAIN-E57B1/6; TISSUE=Intestine; Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.; Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases. |
| RA | Whiting M., Maddan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., |
| DR | EMBL; AF176841; AAD56003.1; -. |
| DR | AF176840; AAD56002.1; -. |
| Query Match | Score 1708; DB 2; Length 395; |
| Best Local Similarity | 83.1%; Pred. No. 4.7e-16; |
| Matches | 328; Conservative 24; Mismatches 43; Indels 0; Gaps 0; |
| Qy | 1 MWLPRVSSTAVTALLAQ-TFLILFLVSRGPSSPAGGEARHFLVLSSWRGSSFWQL 60 |
| Db | 1 MWLPRFSSTVMULLMWYQTGLVFLSYRQPSSPAGLGERVHLVLSWRGSSFWQL 60 |
| Qy | 61 NORPDVYLMPEPAWHVWTLSQGSAAATHMAYRDLMRFDAYLPWRRNLSD 120 |
| Db | 61 SORPDVYLMPEPAWHVWTLSQGSAAATHMAYRDLMRFDAYLPWRRNLSD 120 |
| Qy | 121 FQWAVSRLCSPPACSAFPGRGAISSCAEVCKPLCARQSTFLAREACRSYSHVILKEVRFN 180 |
| Db | 121 FQWAVSRLCSPPACSAFPGRGTSKQDVCTKLTQFSLAREACRSYSHVILKEVRFN 180 |
| Qy | 181 LQVLYPLISDPALNLRVHLVEDPRAVLSREOTAKALARNDIVLGNTGWAEADPLR 240 |
| Db | 181 LQVLYPLISDPALNLRVHLVRDPRAVLSREOTAKALARNDIVLGNTGWAEADPLR 240 |
| Qy | 241 YREVCRSHVRLAEEATLKPPPLGRYRVRFEDLAREPLAEIRALYAFTGLSLTPQL 300 |
| Db | 241 WNEVCRSHVRLAEEAHLKPPPLQDYLRLVRYTFLPTVIRETAFTGLSLTPQL 300 |
| Qy | 301 AWHNITHSGGPGRARRAKFTSSRNALNSQMRHALPKAKRVOELCAGALQLGYR 360 |
| Db | 301 TWHNITHSGGPGRARRAKFTSSRNALNSQMRHALPKAKRVOELCAGALQLGYR 360 |
| Qy | 361 VVSDEORNLADLVLPGLNGTWASSTASHPRN 395 |
| Db | 361 VHSELEQRDSLDDLRGMDSPKWAASSTEKOPES 395 |
| RESULT 5 | |
| ID | Q6GN39 |
| AC | Q6GN39; |
| DT | 05-JUL-2004 (TREMBLrel. 27, Created) |
| DT | 05-JUL-2004 (TREMBLrel. 27, Last sequence update) |
| DB | MGCB31048 protein. |
| GN | Name=MGCB31048; |
| OS | Xenopus laevis (African clawed frog). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Amphibia; Batracia; Anura; Mesobatrachia; Pipoidea; Pipidae; |
| OC | Xenopoda; Xenopus. |
| NCBI_TaxID | 6355; |
| [1] | |
| RN | SEQUENCE FROM N.A. |
| RC | TISSUE=Embryo; RCMLINE=2238257; PubMed=10.1073/pnas.242603899; |
| RX | DOI=10.1073/pnas.242603899; |
| RA | Straubberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D., Altschul S.F., Zeeberg B.R., Buetow K.H., Schaeffer C.F., Bhat N.K., Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Haile F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Tohuyuki S., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muñiz D.M., Sodergren B.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Maddan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., |
| RA | Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases. |
| RA | EMBL; AF176841; AAD56003.1; -. |
| DR | AF176840; AAD56002.1; -. |

| | | |
|-----------------------|--|--|
| RA | Krzywinski M.I., Skalska U., Smailus D.E., Schein J.B., Jones S.J., Marra M.A.; "Generation and analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). | RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K., Masuho Y., Ono T., Okano K., Yoshihara Y., Aotsuka S., Sasaki N., Hattori A., Okumura K., Iwayanagi T., Niromiya K., Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases. |
| RT | SEQUENCE FROM N.A. | DR EMBL; AK074746; BAC11177_1; -. |
| RP | TISSUE=embryo; | DR Genew; HGNC:1972; CHST4. |
| RC | RC PubMed:2341132; PubMed:12454917; DOI:10.1002/dvdy.10174; Klein S.L., Straubberg R.L., Wagner L., Pontius J., Clifton S.W., Richardson P.; "Generic and genomic tools for Xenopus research: The NIH Xenopus Initiative"; Dev. Dyn. 225:384-391 (2002). | DR GO; GO:0008146; F:sulfotransferase activity; IEA. |
| RA | RA TISSUE=embryo; | DR GO:0016740; F:transferase activity; IEA. |
| RL | RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases. | DR InterPro; IPR000863; Sulfotransfase. |
| DR | DR GO: GO:0008146; F:sulfotransferase activity; IEA. | DR Pfam; PF00885; Sulfotransferase_1. |
| DR | DR SEQUENCE 392 AA; 45969 MW; | KW Transferase. |
| SQ | SQ SEQUENCE 386 AA; 45160 MW; 861869348319E42A CRC64; | SQ SEQUENCE 386 AA; 45160 MW; 861869348319E42A CRC64; |
| Query Match | | |
| Qy | Qy 14 LLLAQTFULLFLYFSRPL--GPSSPAGEGEARVHVLVLSWSRGSSVYFSQL 59 | Qy 14 LLLAQTFULLFLYFSRPL--GPSSPAGEGEARVHVLVLSWSRGSSVYFSQL 59 |
| Db | Db 2 LJPKKMFLVSYQVAILALFFHMYSHNISLMSMKQPERPRAHVLYLSWSGSFSYFSQL 61 | Db 2 LJPKKMFLVSYQVAILALFFHMYSHNISLMSMKQPERPRAHVLYLSWSGSFSYFSQL 61 |
| Qy | Qy 60 FNQHPDVFYLMEPAWHWTUQSQGSAITLAVARDLVPSFLCDMDVFDAYL-PWRENLIS 118 | Qy 60 FNQHPDVFYLMEPAWHWTUQSQGSAITLAVARDLVPSFLCDMDVFDAYL-PWRENLIS 118 |
| Db | Db 62 FCQHPDVFYLMEPAWHWTMTPQOSTAWMLHMAVRDLIRAVEFLCDMSFDAYMEPGPRRQS 121 | Db 62 FCQHPDVFYLMEPAWHWTMTPQOSTAWMLHMAVRDLIRAVEFLCDMSFDAYMEPGPRRQS 121 |
| Best Local Similarity | | |
| Matches | Matches 220; Conservative 68; Mismatches 91; Indels 4; Gaps 3; | Qy 119 DLFWQAYWSRALCSPCPASAFPQGAISSEAVCKPLCARQSFTLAREACRSYSHVVLKEVRF 178 |
| Db | Db 122 SFLQWENSRALCSAPAD1IPDEIETRAHCRLLCSQPFEVVEACRSYSHVVLKEVRF 181 | Db 122 SFLQWENSRALCSAPAD1IPDEIETRAHCRLLCSQPFEVVEACRSYSHVVLKEVRF 181 |
| Qy | Qy 179 FNQLQVYPLLSDPALNLRIVHLVYRDPAVLRSRECTAKALARNDTGTLGTGNTGW-BADP 237 | Qy 179 FNQLQVYPLLSDPALNLRIVHLVYRDPAVLRSRECTAKALARNDTGTLGTGNTGW-BADP 237 |
| Db | Db 182 FNQSLSPFLPKDPSLNTHVHLVYRDPAVFSRERTGDLMTDNRTMGHQEQLKEDQ 241 | Db 182 FNQSLSPFLPKDPSLNTHVHLVYRDPAVFSRERTGDLMTDNRTMGHQEQLKEDQ 241 |
| Qy | Qy 238 GLRVVPRVCRSHVRIAAATLKP-PPPLRGRYRLVRFEDLAREPLAIRALYAFTGJSLT 296 | Qy 238 GLRVVPRVCRSHVRIAAATLKP-PPPLRGRYRLVRFEDLAREPLAIRALYAFTGJSLT 296 |
| Db | Db 242 PYVMQVICQSOLEIYK--TIOSLPKRQERYLILRYEDLARAPVACTSRSRMYEFVGJBLF 299 | Db 242 PYVMQVICQSOLEIYK--TIOSLPKRQERYLILRYEDLARAPVACTSRSRMYEFVGJBLF 299 |
| Qy | Qy 297 POLEANHTNITHGSGPQARREAFKTSRNALNSQAWRHALPFAKIIRRQELCAGAL-QLL 356 | Qy 297 POLEANHTNITHGSGPQARREAFKTSRNALNSQAWRHALPFAKIIRRQELCAGAL-QLL 356 |
| Db | Db 300 PHQQTWHNITRPGKGM--DHAFTHINRDAINVSQAWRSLPYEKVSRQLKACGDAMNL 357 | Db 300 PHQQTWHNITRPGKGM--DHAFTHINRDAINVSQAWRSLPYEKVSRQLKACGDAMNL 357 |
| Qy | Qy 357 GYRIVYSEDEQPNLA-DLVLPRLGLNGTWASSTASH 392 | Qy 357 GYRIVYSEDEQPNLA-DLVLPRLGLNGTWASSTASH 392 |
| Db | Db 358 GYRHVREBQEQQNLQNLQNL-----STWTYPEQIH 386 | Db 358 GYRHVREBQEQQNLQNLQNL-----STWTYPEQIH 386 |
| RESULT 7 | | |
| Q9Y5R3 | Q9Y5R3 PRELIMINARY; PRT; 386 AA. | Q9Y5R3 PRELIMINARY; PRT; 386 AA. |
| AC | AC O9Y5R3 | AC O9Y5R3 |
| RN | RN SEQUENCE FROM N.A. | RN SEQUENCE FROM N.A. |
| RC | RC TISSUE=Tonsil; | RC TISSUE=Tonsil; |
| RX | RX MEDLINE=99-64336; PubMed=10330415; DOI=10.1089/jcb.145.4.899; | RX MEDLINE=99-64336; PubMed=10330415; DOI=10.1089/jcb.145.4.899; |
| RA | RA Bistrik A., Bhakta S., Lee J.X., Belov Y.Y., Gunn M.D., Zuo F.R., Huang C.C., Kannagi R., Rosen S.D., Hemmerich S.; | RA Bistrik A., Bhakta S., Lee J.X., Belov Y.Y., Gunn M.D., Zuo F.R., Huang C.C., Kannagi R., Rosen S.D., Hemmerich S.; |
| RT | RT "Sulfotransferases of two specificities function in the reconstitution of high endothelial cell ligands for L-selectin."; | RT "Sulfotransferases of two specificities function in the reconstitution of high endothelial cell ligands for L-selectin."; |
| GN | GN Name=GST3; | GN Name=GST3; |
| OS | OS Homo sapiens (Human) | OS Homo sapiens (Human) |
| OC | OC Mammalia; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Primates; Catarrhini; Hominidae; Homo. | OC Mammalia; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Primates; Catarrhini; Hominidae; Homo. |
| RN | RN [1] | RN [1] |
| RP | RP SEQUENCE FROM N.A. | RP SEQUENCE FROM N.A. |
| RC | RC TISSUE=Tonsil; | RC TISSUE=Tonsil; |
| DT | DT 01-OCT-2002 (TREMBLrel. 22, Created) | DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update) |
| DT | DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update) | DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update) |
| DE | DE Hypothetical protein FLJ90265. | DE Hypothetical protein FLJ90265. |
| OS | OS Homo sapiens (Human). | OS Homo sapiens (Human). |
| OC | OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| OX | OX NCBI_TaxID=9606; | OX NCBI_TaxID=9606; |
| RN | RN [2] | RN [2] |
| RP | RP SEQUENCE FROM N.A. | RP SEQUENCE FROM N.A. |
| RC | RC TISSUE=Tonsil; | RC TISSUE=Tonsil; |

| | | |
|-----------------------|---|---|
| RX | MEDLINE:21096027; PubMed=11181564; DOI=10.1093/glycob/11.1.75; | NCBI_TaxID=9606; |
| RA | Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R., | [1] SEQUENCE FROM N.A. |
| RT | "Chromosomal localization and genomic organization of the galactose/N-acetylgucosamine/N-acetylglucosamine 6-O-sulfotransferase gene family"; | TISSUE:Brain; |
| RT | Glycobiology 11:75-87(2001). | MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; |
| RL | SEQUENCE FROM N.A. | RC |
| RN | SEQUENCE FROM N.A., PubMed=11439191; DOI=10.1016/S0092-8674(01)00394-4; | RN TISSUE:Brain; |
| RX | Yeh J.-C., Hiraoka N., Petryaki B., Nakayama J., Ellies L.G., | RX Straubberg R.L., Feingold B.A., Grouse L.H., Derge J.G., Schuler G.D., |
| RA | Rabuka D., Hindsgaul O., Marth J.D., Lowe J.B., Fukuda M.; | RA Klauser R.D., Collins F.S., Wagner L., Shennan C.M., Schaefer C.P., Bhat N.K., |
| RT | "Novel sulfated lymphocyte homing receptors and their control by a core extension beta 1,3-N-acetylglucosaminyltransferase."; | RA Altshul S.P., Jordan H., Moore T., Max S.I., Wang J., Heich P., |
| RT | Cell 105:957-969 (2001). | RA Hopkins R.P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., |
| RL | EMBL; AF131225; ADD32015.1; | RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., |
| DR | EMBL; AF148177; AAG4816.1; | RA Brownstein M.J., Usdin T.B., Toshiyuki S., Canninci P., Prange C., |
| DR | GO: 0008146; P:sulfotransferase activity; TAS. | RA Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., |
| DR | GO: 0007155; P:cell adhesion; TAS. | RA Bosak S.A., McEvani P.J., McKernan K.J., Malek J.A., Gunnarathne P.H., |
| DR | GO: 0006928; P:cell motility; TAS. | RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., |
| DR | GO: 0007267; P:cell-cell signaling; TAS. | RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., |
| DR | GO: 0006935; P:immune response; TAS. | RA Fahey J., Helton B., Kettman M., Madan A., Rodrigues S., Sanchez A., |
| DR | GO: 0006477; P:protein amino acid sulfation; TAS. | RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., |
| DR | InterPro: IPR000863; Sulfotransf erase_1. | RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Krzywinski M.I., Grimwood J., Myers R.M., Butterfield Y.S., |
| DR | InterPro: IPR000863; Sulfotransf erase.. | RA Jones S.J., Maria M.A., Smaila D.E., Schein J.E., Proc. Natl. Acad. Sci. U.S.A. 99:16699-16703 (2002). |
| RN | [2] | RN |
| RP | SEQUENCE FROM N.A. | RP |
| RC | TISSUE:Brain; | RC |
| RA | Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases. | RA Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases. |
| DR | EMBL; BC035282; AAH35282.1; | DR EMBL; BC035282; AAH35282.1; |
| DR | GO: GO:0008146; P:sulfotransferase activity; IEA. | DR GO: GO:0008146; P:sulfotransferase activity; IEA. |
| DR | InterPro: IPR000863; Sulfotransf erase_1. | DR InterPro: IPR000863; Sulfotransf erase_1. |
| DR | Pfam: PF00655; Sulfotransf erase_1. | DR Pfam: PF00655; Sulfotransf erase_1. |
| DR | SEQUENCE 370 AA; EA8D76EA4E73C625 CRC64; | DR SEQUENCE 370 AA; EA8D76EA4E73C625 CRC64; |
| Query Match | 49.6%; Score 1019; DB 2; Length 386; | Query Match 49.2%; Score 1012; DB 2; Length 370; |
| Best Local Similarity | 54.5%; Pred. No. 9.7e-78; | Best Local Similarity 57.9%; Pred. No. 3.6e-79; |
| Matches | 216; Conservative 45; Mismatches 107; Indels 28; Gaps 7; | Matches 206; Conservative 43; Mismatches 93; Indels 14; Gaps 6; |
| Qy | 14 LLLQTFLPLFLVSPR-----GPPSPAGGEARVAVYLVLSSWRSSESSFYVQL 59 | Qy 40 RVEFLVLSWRSSSSFYVQLFNLQHPDFYLMERAWHWTTLSQGSAAATHMAYVRLYRVS 99 |
| Db | 2 LLPKMKMLFLVPSQMATALLFFHMYSHNISLSMSMAQPERMHLVVLSSWRSSESSFYVQL 61 | Db 26 RMHVLVLSWRSSSSFYVQLFGHDFYLMEPWHTMTEFKQSTANMLHMAVRLYRVA 85 |
| Qy | 60 ENQHPDVFLVMEPAWHWTTLSQGSAAATHMAYVRLVSRALCSPACDIPDEIIPRACRLLS 118 | Qy 100 FLCDMDVFDAYL-PWRLNLSDFQWVSRALCSPACDIPDEIIPRACRLLS 158 |
| Db | 62 FGQHPDVFLVMEPAWHWTTLSQGSAAATHMAYVRLVSRALCSPACDIPDEIIPRACRLLS 121 | Db 86 FLCDMSVFDAYMGPGRBQSSLFQWVSRALCSPACDIPDEIIPRACRLLCSQQPF 145 |
| Qy | 119 DLFWAVASRALCSPACSAFPGRASISSEACKPLCAROSFTLAREACRSYSHVYLKEYRF 178 | Qy 159 TLAREACRSYSHVYLKEYRFENLQVLYPLSLSDPANLRLVHLDPRAYLRSREQTAKAL 218 |
| Db | 122 SLFWENSRALCSPACDIPDEIIPRACRLLCSQQPFEEVKACRSYSHVYLKEYRF 181 | Db 146 EVYKEACRSYSHVYLKEYRFENLQVLYPLSLSDPANLRLVHLDPRAYLRSREQTAKAL 205 |
| Qy | 179 FNQVLYPLISDPALNLRVHLYDPRAVLRSRQTAKLARDNGIVLGTGTVW-EADP 237 | Qy 219 ARDGIVLVTGTVW-EADPGLVRVRECVRSYHIAEATLKP-PFPLGRYRLVRFEDL 276 |
| Db | 182 FNQSLVPLKJDPSNLH1VHLYDPRAVFRSRRTKGDLMIDSRYVNGQEQLKCKEDQ 241 | Db 206 MDSRIVNGQEQLKCKEDQPYVYMMQVICQSLEYK--TIQSLPKALQERYLVRYEDL 263 |
| Qy | 238 GLRVYREVCRSHVYIAEATLKP -PFPFLGRYRLVRFEDAREPLAEITALYAGLSLT 296 | Qy 277 AREPLAEITALYAGLSLTQPLQBAWHTHNTGSGPGRARREAPKTSNANLYSQWHRHA 336 |
| Db | 242 PYMVQVICQSLEYK--TQSLSPKALQERYLLVRYEDLARAVTAQTSRMYEVGLEFL 299 | Db 264 ARAPVQATSRMVFYVGELPLPHQIOTWANTIRGKGMG--DHAFHTNARDALNYSQWHRKA 321 |
| Qy | 297 PQLAEWHTHNTGSGPGRARREAPKTSNANLYSQWHRHALPFAKIRYQELCAGALOLL 356 | Qy 337 LPFAKIRYQEVQELCAGALOLLGYPRVYSEDEORNIALDLYLPRGLNGFTWASSSTASH 392 |
| Db | 300 PHQQTWVHNTRGKGMG--DHAFHTNARDALNYSQWHRSLPYSKVRSLQKACGDMNL 357 | Db 322 LPYEVKSILPKACDANLGYHVRSLQEQRNLLDDL-----STWTVPOQIH 370 |
| RESULT 8 | | |
| ID | Q8IVY46 | PRELIMINARY; PRT; 370 AA. |
| AC | Q8IVY46; | |
| DT | 01-MAR-2003 (TREMBrel. 23, Created) | Q9R111 PRELIMINARY; PRT; 388 AA. |
| DT | 01-MAR-2003 (TREMBrel. 23, Last sequence update) | AC Q9R111; |
| DT | 01-MAR-2004 (TREMBrel. 26, Last annotation update) | DT 01-MAY-2000 (TREMBrel. 13, Created) |
| DB | CHST4 protein. | DT 01-MAY-2000 (TREMBrel. 13, Last sequence update) |
| OS | Homo sapiens (Human). | DT 05-JUL-2004 (TREMBrel. 27, Last annotation update) |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates. | DE L-selectin ligand sulfotransf erase (Chst4 protein). |

| | | | |
|--|---|--|--|
| | | 231 | HLETIKEEDOPYYAMKILCKSQDIVAKATQL- PEALQORYLFYRVEDLVRPLAQTRR 288 |
| Mus musculus (Mouse) | S | SEQUENCE FROM N.A. | |
| Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | C | STRAIN=FVB/N; TISSUE=Colon; | |
| Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | C | MEDLINE=9361534; PubMed=10435581; DOI=10.1016/S1074-7613(00)80083-7; | Qy |
| [1] _TAXID=10090; | N | Kraoaka N., Petryniak B., Nakayama J., Tsuibo S., Suzuki M., Yeh J.C., Izawa D., Tanaka T., Miyasaka M., Lowe J.B., Fukuda M., "A novel, high endothelial venule-specific sulfotransferase expresses 6-sialyl Lewis(x), an Ig-selectin ligand displayed by CD34.", Immunity 11:79-89 (1999). | Db |
| [2] | N | SEQUENCE FROM N.A. | |
| | C | STRAIN=FVB/N; TISSUE=Colon; | P |
| | C | MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; | X |
| | A | Strausberg R.L., Ringold F.A., Grouse L.H., Dege J.G., Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schulter G.D., Altchul S.F., Zeeberg B., Buetow K.H., Bhattacharjee A., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marsusina K., Farmer A.A., Rubin G.M., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshimatsu S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulilahy S.J., Bosack S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzyz D.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting M., Madan A., Keittman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schneich A., Schein J.E., Jones S.M., Marra M.A., "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). | Qy |
| [3] | N | SEQUENCE FROM N.A. | P |
| | C | STRAIN=FVB/N; TISSUE=Colon; | C |
| | A | Submitted (SEP 2003) to the EMBL/GenBank/DBJ databases. | C |
| | A | EMBL: BC57886; AA05579.1; -. | C |
| | R | MGD: MGI:139479; Chst4. | R |
| | R | GO: GO:0016021; C:integral to membrane; TAS. | R |
| | R | InterPro: IPR000863; Sulfotransferase. | R |
| | R | Pfam: PF00685; Sulfotransferase_1. | R |
| | R | Selectin; Transferease. | R |
| | Q | SEQUENCE 388 AA; 44635 MW; 6D5371AFB6884AEE CRC64; | Q |
| | Y | Query Match 47.9%; Score 984; DB 2; Length 388; Best Local Similarity 53.5%; Pred. No. 9e-75; Gaps 7; Matches 200; Conservative 42; Mismatches 97; Indels 42; Gaps 7; | Y |
| b | Y | 14 LLLAGTEFLFLVS-----2 MLLXGRMLGSGQVIVVAlFIRMSVHRLSQRERESRERQ | Y |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRALCSAVCDPPANEISSPKCKLICQQDPMDVKCRSH 170 169 SHVVLKEYRFENLQVLYPLLSDPAINLRVTHYLDPRAVIRSREQATAKALARDNGIVGT 228 171 GFVWLKEYRSLQALYPLTDPSLNLYHVLVDRPRAYERSRHTTBLVYDHSVHQ 230 229 NGTVEADGLRVVREVCSHVIAEA-ATLKKPPFLNQGRILVRFELARPLAETRA 286 | b |
| b | Y | 50 RSGSEFVGQENQHDVFVIMERPAWVWITLQSSESAATIMAVDVLRSVFLCDMDVDA 109 51 RSGSEFVGQFLGQHPDVFLMEPAWVWMTTSSTAWKLHMAVDLRLSFLCDMSVFD 110 110 YL-PWRRNLSDLFOWAVSRALCSPACASAPRGTAISSEACKPCLCARQSFETLREACSY 168 111 YMNPQPKRQSSLQFEQSRAL | |

| | |
|------------|--|
| DE | N-acetylglucosamine-6-O-sulfotransferase long form. |
| GN | Name=Chst2; |
| OS | Mus musculus (Mouse). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| NCBI_TaxID | 10090; |
| [1] | |
| RN | SEQUENCE FROM N.A. |
| RP | RC TISSUE:Whole embryo; |
| RX | MEDLINE=9830482; PubMed=9712885; DOI=10.1074/jbc.273.35.22577; |
| RA | Uchimura K., Muramatsu H., Kadomatsu K., Fan Q., Kurokawa N., |
| RA | Matsuoka C., Kanagai R., Habuchi O., Muramatsu T., |
| BA | "Molecular cloning and characterization of an N-acetylglucosamine-6-O- |
| RT | sulfotransferase," |
| RT | J. Biol. Chem. 273:22577-22583 (1998). |
| RL | DR ENSEMBL; AB011452; BAA32138.1; -. |
| DR | DR MGB; MG1:1831160; Chs22. |
| DR | GO; GO:000515; C:extracellular space; TAS. |
| DR | GO; GO:0008446; F:sulfotransferase activity; IDA. |
| DR | GO; GO:0005644; P:N-acetylglucosamine metabolism; TAS. |
| DR | InterPro; IPR000863; Sulfotransferase. |
| DR | Pfam; PF00655; Sulfotransfer_1; 1. |
| KW | |
| SQ | SEQUENCE 530 AA; 57814 MW; A113E1B73C363BC CRC64; |
| | Query Match 3.24%; Score 663; DB 2; Length 530; |
| | Best Local Similarity 3.9-5.1; Pred. No. 2e-47; Length 530; |
| | Matches 146; Conservative 67; Mismatches 123; Indels 34; Gaps 9 |
| Oy | 36 GGEARHVVLVSSWRSGSSFSYQGLFNOHPDVYLMPEAHVWTRTSGSAATLHIMAVRDL 95 |
| Db | 160 GGDRQLYVFVFTWRSGSSFFGELFVNPEVFYLTPVHWWQKLYPGDVAISLGQAAIRDM 21.9 |
| Qy | 96 VRSVFLCDMDVFDAYLP--WRRNLSDL--PQWAVSRLACSPCPACSAFPRGAIIS--SEAV 148 |
| Db | 220 LSALYRCPLSVQPLQYSPAGSGCRNLITLGFEAGATKVVKVVSPCPAYKEVYGLDDRV 27.9 |
| Qy | 149 CKELCARQSFATLAREACRSYSHVVKLEVRFNQLQVYPLISDPALNRLQVHLVDFDPRAVL 208 |
| Db | 280 CRK-CPPORLARFEECRKYTVIIGVYRVDVAVLAFLKDVKIHLVDFDPRAVA 33.8 |
| Qy | 209 RSEQQTAKALARDN-----GIVLGT--NGTWVEAD-PGIVRVREV 245 |
| Db | 339 SSIRISSHGLIRESLQVRSRDPRAHMPFLEAGHKLGAKEGGPADDYHALGAMEVI 398 |
| Qy | 246 CRSHVRIAEATLKPPEFLGRYLVRFEDLAREPLAEITALYAFGLSLTPOQEAWIHN 305 |
| Db | 399 CNSMAKLTQ-LQPPMWLQGHLYVRYEDLGVDPVKTLRVYDGLAVSPENHQFAIN 457 |
| Qy | 306 ITFHSGPCARRPAFKTSSRNALNVSOAHRHALPAKIRYQELCAGALOLQGYPVYSED 365 |
| Db | 458 MTSGSSSSK--PFVVSARNATAANAWRALTFFQQIKQVBEFFCYQPMAVLGYERVNSPE 515 |
| Qy | 366 EQRNLLADLV 375 |
| Db | 516 EVRDLSKTL 525 |
| | RESULT 13 |
| QBWV3 | PRELIMINARY; PRT; 530 AA. |
| AC | QBWV3; |
| DT | 01-JUN-2003 (TREMBLrel. 24, Created) |
| DT | 01-JUN-2003 (TREMBLrel. 24, Last sequence update) |
| DT | 01-MAR-2004 (TREMBLrel. 26, Last annotation update) |
| DE | Chst2 protein. |
| GN | Name=Chst2; |
| OS | Mus musculus (Mouse). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| NCBI_TaxID | 10090; |
| RN | SEQUENCE FROM N.A. |

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 23, Last annotation update)
 DE N-acetylglucosamine-6-O-sulfotransferase (GlcNAc6ST).
 GN Name=GN6ST;
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI TaxID=9606;
 RN [1] -
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 MEDLINE=98391845; PubMed=9722682;
 RX RA Mitsuoka C., Muramatsu K., Kaname T., Ogawa H., Yamakawa T., Fan Q., Ozaki T., Nakagawa A., Kadomatsu K., Muramatsu T.;
 RT "Human N-acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of 6-sulfo sialyl Lewis X: molecular cloning, chromosomal mapping, and expression in various organs and tumor cells.";
 RL J. Biochem. 124:670-678(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pleacenta;
 MEDLINE=91042394; PubMed=11042394; DOI=10.1016/S0303-4165(00)00136-7;
 RA Sakaguchi H., Kitagawa H., Sugahara K.;
 RT "Functional expression and genomic structure of human N-acetylglucosamine-6-O-sulfotransferase that transfers sulfate to b-N-acetylglucosamine at the nonreducing end of an N-acetyllactosamine sequence";
 RL Biochim. Biophys. Acta 1523:269-276 (2000).
 DR EMBL; AB014679; BA034265; 2;
 DR EMBL; AB021124; BAB16886; 1;
 DR EMBL; AB021125; BAB16887; 1;
 DR GO; GO:0008116; F:sulfotransferase activity; IEA.
 DR GO; GO:0016740; F:transf erase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase.
 DR Pfam; PF00685; Sulfotransfer_1; 1.
 KW Transferase.
 SQ SEQUENCE 483 AA;
 SEQUENCE 52787 MN;
 SEQUENCE 82779 TD44931BA18 CRC64;
 Query Match Score 657.5; DB 2; Length 483;
 Best Local Similarity 38.2%; Pred. No. 5.2e-47;
 Matches 147; Conservative 70; Mismatches 125; Indels 43; Gaps 10;

YQ 30 GPSSPAG-----GEARHVLVLLSSWRSRSSFVGOLFNOHDVFYLMEPAWHWNTL 80
 Db 98 GVAAPGNGTRGTGVGDKRQLVYFTWRSGSSFFEGELNQNDBVFPEPYWHWQKL 157

QY 81 SGQSATLHMARDLVRSVFLCDMDVEDAYLP--WRRNLSDL--FQWAVSRLCSPPAC 135
 Db 158 YPGDRAVSQCGARDMLSAVRCDASVFLQXSPAGSGGRNLTIGFGARTNKVYCSPPC 217

QY 136 SAFRGAIS--SEAVCKPLCARQSTLARECRSYHVVLYKEVRFNOLYPLPSDPAL 193
 Db 218 PAYRKEVVGVLDDRYCKK-CPPQRLLARFEEBECRKYRTLVIGVRFEDAVALPLRDPAL 276

QY 194 NLRIVHLVRDPRAVLRSRECTAKALARDN-----GIVLGTT--NGT 231
 Db 277 DLKVTHLVRDPRAVASSRIRSRHGLIRESIQVVERSDPRAHMPLEAIGHKGAKKEGV 336

QY 232 WFEAD-PGLFVVRREVCRSHVRIAETLKPPLFLRGYRLVRFEDLAREPLAEIRALYAF 290
 Db 337 GGPADYHALGAMEVICNSMKTLOTA-LQPDWLGQHYLIVVRYDLGPVKTLRYYDF 395

QY 291 TGLSLTPQEAWHNITHGGPGRARAFFTSNNANLYSQMRHALPPAKIRVQELCA 350
 Db 396 VGLLVSPEMFQALNMTCGSSSSK--PFVVSARNATOAAANARTALTFOQIKOVEEFCY 453

QY 351 GALQJGYRPVYSEDEQRNIALDLY 375
 Db 454 QPMAVLGJYERVNNSPBBVKDLSKTL 478

RESULT 15
 ID Q9Y4C5 PRELIMINARY;
 AC Q9Y4C5_QRGZNS; Q9Y6FP;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DB Long form of N-acetylglucosamine-6-O-sulfotransferase (GlcNAc6ST)
 DE (Carbohydrate sulfotransferase 2).
 GN Name=GN6ST; Synonyms=CHST2;
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 MEDLINE=98391845; PubMed=9722682;
 RX RA Uchimura K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q., Mitsuoka C., Kannagi R., Habuchi O., Yokoyama I., Yamamura K., Muramatsu T.;
 RA Ozaki T., Nakagawa A., Kadomatsu K., Muramatsu T.;
 RT "Human N-acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of 6-sulfo sialyl Lewis X: molecular cloning, chromosomal mapping, and expression in various organs and tumor cells.";
 RL J. Biochem. 124:670-678(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Umbilical vein endothelium;
 MEDLINE=99168906; PubMed=1049551; DOI=10.1006/geno.1998.5653;
 RA Li X., Tedder T.F.;
 RT "CHST1 and CHST2 sulfotransferases expressed by human vascular endothelial cells: cDNA cloning, expression, and chromosomal localization.";
 RL Genomics 55:345-347(1999).
 DR EMBL; AB014680; BAA34266; 2;
 DR EMBL; AF083066; AAD20881; 1.;
 DR Genew; HGNC:1970; CHST2.
 DR GO; GO:0008116; F:sulfotransferase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR InterPro; IPR000863; Sulfotransferase_1.
 KW Transf erase.
 SQ SEQUENCE 530 AA; MW: A82CA227B9D5651B CRC64;
 Query Match Score 657.5; DB 2; Length 530;
 Best Local Similarity 38.2%; Pred. No. 5.9e-47;
 Matches 147; Conservative 70; Mismatches 125; Indels 43; Gaps 10;

QY 30 GPSSPAG-----GEARHVLVLLSSWRSRSSFVGOLFNOHDVFYLMEPAWHWNTL 80
 Db 145 GVAAPGNGTRGTGVGDKRQLVYFTWRSGSSFFEGELNQNDBVFPEPYWHWQKL 204

QY 81 SGQSATLHMARDLVRSVFLCDMDVEDAYLP--WRRNLSDL--FQWAVSRLCSPPAC 135
 Db 205 YPGDRAVSQCGARDMLSAVRCDASVFLQXSPAGSGGRNLTIGFGARTNKVYCSPPC 264

QY 136 SAFRGAIS--SEAVCKPLCARQSTLARECRSYHVVLYKEVRFNOLYPLPSDPAL 193
 Db 136 SAFRGAIS--SEAVCKPLCARQSTLARECRSYHVVLYKEVRFNOLYPLPSDPAL 193

QY 194 NLRIVHLVRDPRAVLRSRECTAKALARDN-----GIVLGTT--NGT 231
 Db 265 PAYRKEVVGVLDDRYCKK-CPPQRLLARFEEBECRKYRTLVIGVRFEDAVALPLRDPAL 323

QY 291 TGLSLTPQEAWHNITHGGPGRARAFFTSNNANLYSQMRHALPPAKIRVQELCA 350
 Db 324 DLKVTHLVRDPRAVASSRIRSRLGILRESLQVVRSDPRAHMPLEAIGHKGAKKEGV 383

QY 384 GGPADYHALGAMEVICNSMKTLOTA-LQPDWLGQHYLIVVRYDLGPVKTLRYYDF 442

QY 291 TGLSLTPQEAWHNITHGGPGRARAFFTSNNANLYSQMRHALPPAKIRVQELCA 350
 Db 443 VGLLVSPEMFQALNMTCGSSSSK--PFVVSARNATOAAANARTALTFOQIKOVEEFCY 500

QY 351 GALQJGYRPVYSEDEQRNIALDLY 375

Db 501 QPMAVLGVERNSPEEVXDLSKTLL 525

Search completed: June 23, 2005, 08:49:01
Job time : 44.3472 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 08:32:43 ; Search time 132.064 Seconds
(without alignments)
3578.727 Million cell updates/sec

Title: US-10-697-828-15
Perfect score: 6532
Sequence: 1 MPKGAPPWIMALMFTGHLL.....LIENICWTLMDRGYPKFMD 1222

Scoring table: BLOSUM62
Gapop 10.0 , Gapsct 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqD19808:*
- 2: geneseqD19904:*
- 3: geneseqD20008:*
- 4: geneseqD20018:*
- 5: geneseqD20024:*
- 6: geneseqD20038:*
- 7: geneseqD20038b:*
- 8: geneseqD20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Match Length | DB ID | Description |
|------------|-------|-------|--------------|------------|---|
| 1 | 6532 | 100.0 | 1222 | 4 PAY72642 | Aay72642 Human gly Aab1149 Human sec Aae33541 Human nov |
| 2 | 6509 | 99.6 | 1222 | 6 AAE1849 | Aay72643 Mouse gly Aae33542 Human nov |
| 3 | 6465 | 99.0 | 1212 | 6 AAE33541 | Aab33735 Human pro Adp56673 Human pro Aay72641 Human gly Aab9410 Human pro |
| 4 | 5899 | 90.3 | 1207 | 4 PAY72643 | Aay72643 Mouse gly Aae33542 Human nov |
| 5 | 5899 | 90.3 | 1207 | 6 AAE33542 | Aab33735 Human pro Adp56673 Human pro Aay72641 Human gly Aab9410 Human pro |
| 6 | 4037 | 61.8 | 755 | 4 AAB93626 | Aay72641 Human gly Aab9410 Human pro |
| 7 | 4037 | 61.8 | 755 | 8 ADR10287 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 8 | 3147 | 48.2 | 596 | 4 AAB94410 | Aay72641 Human gly Aab9410 Human pro |
| 9 | 2597 | 39.8 | 480 | 4 AAB94410 | Aay72641 Human gly Aab9410 Human pro |
| 10 | 1753 | 26.8 | 958 | 3 PAY51120 | Aay72641 Human gly Aab9410 Human pro |
| 11 | 1753 | 26.8 | 958 | 8 ADL83102 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 12 | 1333 | 20.4 | 314 | 4 AAB93626 | Aay72641 Human gly Aab9410 Human pro |
| 13 | 898 | 13.7 | 625 | 8 ADR10287 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 14 | 637 | 9.8 | 128 | 4 ADG27796 | Aay72641 Human gly Aab9410 Human pro |
| 15 | 568 | 8.7 | 125 | 8 FDD12217 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 16 | 568 | 8.7 | 126 | 2 AAY14415 | Aay72641 Human gly Aab9410 Human pro |
| 17 | 431 | 6.6 | 456 | 8 ADM30411 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 18 | 368 | 5.6 | 218 | 4 ADM24176 | Aay72641 Human gly Aab9410 Human pro |
| 19 | 348.5 | 5.3 | 474 | 3 AAB34722 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 20 | 147 | 2.3 | 410 | 7 ADJ6589 | Aay72641 Human gly Aab9410 Human pro |
| 21 | 147 | 2.3 | 411 | 2 AAW61100 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 22 | 147 | 2.3 | 411 | 5 AAE25356 | Aay72641 Human gly Aab9410 Human pro |
| 23 | 147 | 2.3 | 411 | 6 ADU79121 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |
| 24 | 147 | 2.3 | 411 | 7 ADJ3375 | Aay72641 Human gly Aab9410 Human pro |
| 25 | 147 | 2.3 | 411 | 8 ADQ95147 | Aab33826 Human pro Adr10287 Human pro Adg27796 Human nov Adj12217 Human sec Aay14415 Human sec Adp30411 Human PRO Aam24176 Human EST Aab34722 Human sec Adn95522 Human BEC Aaw61100 Keratan s Aae25356 Human cho Abu03503 Angiogeno Adj6589 Human hea Adj6589 Human hea |

ALIGNMENTS

RESULT 1

AAV72642 standard; protein; 1222 AA.

ID AAV72642;

XX AC;

XX DT;

XX 02-MAY-2001 (first entry)

DE Human glycosyl sulfotransferase-6 (GST-6).

XX Human; glycosyl sulfotransferase-6; GST-6; immunosuppressive; therapy;

KW selectin binding inhibitor; gene therapy; inflammation;

KW systemic lupus erythematosus; SLE; rheumatoïd arthritis; diabetes;

KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;

KW glomerulonephritis; myasthenia gravis; Sjögren's syndrome; Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia; denyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis; myocarditis; adult respiratory distress syndrome; eczema; psoriasis; asthma; hypersensitivity; rheumatic fever; tissue rejection.

XX XX Homo sapiens.

OS OS XX Key

FH FH XX Location/Qualifiers

XX FT Domain

XX FT /label= C-terminal_sulfotransferase_domain

XX PN WO200106015-A1.

XX PD 25-JAN-2001.

XX XX (RESC) UNIV CALIFORNIA.

XX PP 19-JUL-2000; 2000WO-US019741.

XX XX PR 20-JUL-1999; 99US-0144694P.

XX PR 13-JUN-2000; 2000US-00593828.

XX DR N-PSDB; AAD02702, AAD02704.

XX PT New Glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications.

XX PS Claim 3; Fig 5B; 128pp; English.

XX CC The present sequence is human glycosyl sulfotransferase-6 (GST-6). GST is

a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non-sulphated selectin ligand, GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in gene inhibiting a selectin mediated binding event. GST is useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarthritis, nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's disease, Grave's disease, adenitis, hypoparathyroidism, pernicious anaemia, demyelinating diseases, cirrhosis, ulcerative colitis, dermatitis, myocarditis, regional enteritis, adult respiratory distress syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation

XX Sequence 1222 AA;

| | | | | |
|--|--|---|--|------|
| CC | Qy | 781 | KFNIAVGLILCISLYLTITPWRFLYSRCLMRWILILVIALWFIELLDWSTCSOPICAK | 840 |
| CC | Db | 781 | KFNIAVGLILCISLYLTITPWRFLYSRCLMRWILILVIALWFIELLDWSTCSOPICAK | 840 |
| CC | Qy | 841 | WTRTEAGSKSKSISSEGHMIDLDPVITSPLGSAEILKOLFENSSDFLYIRVPTAYIDI | 900 |
| CC | Db | 841 | WTRTEAGSKSKSISSEGHMIDLDPVITSPLGSAEILKOLFENSSDFLYIRVPTAYIDI | 900 |
| CC | Qy | 901 | PETELEIDSFDACEWKVSDIRSIGHFRLLRGWLSQDPTKLHQINHLPNPGKLAQY | 960 |
| CC | Db | 901 | PETELEIDSFDACEWKVSDIRSIGHFRLLRGWLSQDPTKLHQINHLPNPGKLAQY | 960 |
| CC | Qy | 961 | FAMNKDKRKFKKRESLPEORSOMKGAFRDAEYTRALRRHLVTPSARPLSISGSWT | 1020 |
| CC | Db | 961 | FAMNKDKRKFKKRESLPEORSOMKGAFRDAEYTRALRRHLVTPSARPLSISGSWT | 1020 |
| XX | Qy | 1021 | LKLHFFQEVLGASMBALIYVDRPAWISYMLYNKSPSLSYLKVNPEHHLAKLFKLEGKGRK | 1080 |
| XX | Db | 1021 | LKLHFFQEVLGASMBALIYVDRPAWISYMLYNKSPSLSYLKVNPEHHLAKLFKLEGKGRK | 1080 |
| Matches 1222; Conservatve 0; Mismatches 0; Indels 0; Gaps 0; | Qy | 1081 | CNLNSGYAFYEPLRKELSKSKSNAVSLSHLWLNLTAAALRINTDLPTSYQLVKFEDI | 1140 |
| Matches 1222; Conservatve 0; Mismatches 0; Indels 0; Gaps 0; | Db | 1081 | CNLNSGYAFYEPLRKELSKSKSNAVSLSHLWLNLTAAALRINTDLPTSYQLVKFEDI | 1140 |
| Qy | 1141 | VHFPORTTERIFAFLGIPLSPASINQILPATSTNLFLYPEGESPTNTNWKQNLPDDE | 1200 | |
| Qy | 1141 | VHFPORTTERIFAFLGIPLSPASINQILPATSTNLFLYPEGESPTNTNWKQNLPDDE | 1200 | |
| Db | 1201 | IKLIEINCMTLMDRGYPKFMD | 1222 | |
| Db | 1201 | IKLIEINCMTLMDRGYPKFMD | 1222 | |
| RESULT 2 | | | | |
| | ABU11849 | | | |
| | ID ABU11849 | standard; protein; 1222 AA. | | |
| | XX | | | |
| | AC ABU11849; | | | |
| | XX | | | |
| | DT 12-FEB-2003 | (first entry) | | |
| | XX | | | |
| | DE Human secreted protein SEC P-4, | INCYTE 3441255CD1. | | |
| | XX | | | |
| | KW Human; SEC P: secreted protein; micro-array; liver disease; hepatitis; | | | |
| | KW cirrhosis; cell proliferative disease; cancer; atherosclerosis; | | | |
| | KW neurological disorder; epilepsy; Huntington's disease; stroke; | | | |
| | KW cardiovascular disorder; hypertension; angina pectoris; allergy; | | | |
| | KW myocardial infarction; immune disorder; inflammatory disorder; AIDS; | | | |
| | KW hypothyroidism; acquired immunodeficiency syndrome; Cushing's syndrome; | | | |
| | KW developmental disorder. | | | |
| | XX | | | |
| | OS Homo sapiens. | | | |
| | XX | | | |
| | PW200286069-A2. | | | |
| | XX | | | |
| | PD 31-OCT-2002. | | | |
| | XX | | | |
| | PF 19-APR-2002; | 2002WO-US012464. | | |
| | XX | | | |
| | PR 20-APR-2001; | 2001US-0285207P. | | |
| | PR 27-APR-2001; | 2001US-0287114P. | | |
| | PR 03-MAY-2001; | 2001US-0289510P. | | |
| | PR 11-MAY-2001; | 2001US-0291814P. | | |
| | PR 18-MAY-2001; | 2001US-0294553P. | | |
| | PR 21-DEC-2001; | 2001US-0343553P. | | |
| | PR 13-FEB-2002; | 2002US-0357002P. | | |
| | PR 20-FEB-2002; | 2002US-0358279P. | | |
| | PR 19-MAR-2002; | 2002US-0366041P. | | |
| | XX | | | |
| | PA (INCY-) INCYTE GENOMICS INC. | | | |
| | XX | | | |
| | PI Klammer AA, Hafalia AJA, Duggan BM, Warren BA, Emerling BM; | | | |
| | Qy | 721 TEHVSVITDYLKTRYENYLGFGGPASVADQQTIRFGLGTQAIYKPVHDRIIPFPFGF | 780 | |
| | Qy | 721 TEHVSVITDYLKTRYENYLGFGGPASVADQQTIRFGLGTQAIYKPVHDRIIPFPFGF | 780 | |
| | Db | 721 TEHVSVITDYLKTRYENYLGFGGPASVADQQTIRFGLGTQAIYKPVHDRIIPFPFGF | 780 | |

| | | | |
|----|---|----|---|
| P1 | Tribouley CM, Aryizu CS, Honchell CD, Nguyen DB, Kallick DA, Yue H; | Qy | 241 ALITGALVTVGDKGSKANIWKQAVDVMKETMPLLNHIVDGSLDEGVAYGSYTAKSVTQY 300 |
| P1 | Au-Young JK, Ramkumar J, Li JX, Thangavelu K, Gietlick DA, Ding Li, | Db | 241 ALITGALVTVGDKGSKANIWKQAVDVMKETMPLLNHIVDGSLDEGVAYGSYTAKSVTQY 300 |
| P1 | Baughn MR, Yao MG, Walia NK, Mason PM, Lal PG, Graul RC, Reddy R; | Qy | 301 VFLAQRFNINNLDNNWILKMHFMFYATLPGFORTVGIADSNNWFGPESQVLFDKPF 360 |
| P1 | Becht SD, Sapperstein SK, Richardson TW, Tran UK, Elliott VS; | Db | 301 VFLAQRFNINNLDNNWILKMHFMFYATLPGFORTVGIADSNNWFGPESQVLFDKPF 360 |
| P1 | Tang YT, Azmazai Y, Yan L, Xu Y; | Qy | 361 ILKNGAGNWLQAQIRKPKDGMVPSTAQRNTLHTXYIWDPQLTQPPOADYGTAKIH 420 |
| XX | DR | Db | 361 ILKNGAGNWLQAQIRKPKDGMVPSTAQRNTLHTXYIWDPQLTQPPOADYGTAKIH 420 |
| XX | PS | Qy | 421 TPNWGUVTYGACLPTONTFSPKGKLGGRAYD1VHFQDYSWIDGWRSENPGEHEP 480 |
| XX | PS | Db | 421 TPNWGUVTYGACLPTONTFSPKGKLGGRAYD1VHFQDYSWIDGWRSENPGEHEP 480 |
| CC | New human secreted proteins (SECP) useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant SECP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, allergies, hepatitis, cirrhosis. | Qy | 481 DQNSFTFAPNGQVFVSSEALYGPKLSHLNVLVFAFPSSQCNKPWEGLGBCAQWLRTG 540 |
| CC | Claim 1; Page 137-140; 192pp; English. | Db | 481 DQNSFTFAPNGQVFVSSEALYGPKLSHLNVLVFAFPSSQCNKPWEGLGBCAQWLRTG 540 |
| CC | XX | Qy | 541 EYGDAAEGLITAQSQHGMVFSGEAISAYSSAMRLSVSYRAUILLNSLTLVVDHIERQ 600 |
| CC | The invention relates to an isolated polypeptide comprising any of 30 naturally occurring amino acid sequences appearing as ABU1846-ABU1175, a naturally occurring amino acid sequence at least 90-98 % identical to the sequences, or a biologically active or immunogenic fragment of the polypeptide. Also included are an isolated polynucleotide encoding the SECP, (including a polynucleotide sequence at least 90-98 % identical to the sequences, their complements, RNA equivalents or fragments comprising at least 60 contiguous nucleotides) a recombinant polynucleotide comprising a promoter sequence operably linked to the SECP polynucleotide, a cell comprising the recombinant polynucleotide, a transgenic organic organism screening for ant/agonists of SECP, generating an expression profile of a sample containing the polynucleotides and an array comprising different nucleotide molecules affixed at distinct physical locations on a solid substrate, where at least one nucleotide molecule comprises a first oligonucleotide or polynucleotide sequence specifically hybridisable with at least 30 contiguous nucleotides of the target polynucleotide. The polypeptides and polynucleotides are useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of SECP, such as liver (e.g. hepatitis, cirrhosis), cell proliferative (e.g. cancer, atherosclerosis), cardiovascular (e.g. epilepsy, Huntington's disease, stroke), myocardial infarction), immune/inflammatory (e.g. acquired immunodeficiency syndrome (AIDS), allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome) disorders (many other diseases and conditions are given in the specification). These are also useful in assessing the effects of exogenous compounds on the target polynucleotide. The SECP or its fragments are useful in screening sequences of SECP. The SECP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. The micro-array is useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. The present sequence is a SECP protein of the invention. | Db | 541 EYGDAAEGLITAQSQHGMVFSGEAISAYSSAMRLSVSYRAUILLNSLTLVVDHIERQ 600 |
| CC | XX | Qy | 601 EDSPINSYSAFFNLDIDFKYIPIKFMRVYNGAMMDIAHYKMFWDHGNSPMASTQE 660 |
| CC | PS | Db | 601 EDSPINSYSAFFNLDIDFKYIPIKFMRVYNGAMMDIAHYKMFWDHGNSPMASTQE 660 |
| CC | XX | Qy | 661 AEQAEEFKCRWTOFVNNTFQMEESTTRIAYVFGPYINVSSCRFIDSSNPGQLQISLNVN 720 |
| CC | PS | Db | 661 AEQAEEFKCRWTOFVNNTFQMEPTTRIAYVFGPYINVSSCRFIDSSNPGQLQISLNVN 720 |
| CC | XX | Qy | 721 TEHVSIIVTDYHNLTKTRENYLGGCFASVADQGQTTRFLGTQAIKVPKVRHDRIIFPGF 780 |
| CC | PS | Db | 721 TEHVSIIVTDYHNLTKTRENYLGGCFASVADQGQTTRFLGTQAIKVPKVRHDRIIFPGF 780 |
| CC | XX | Qy | 781 KENIAVGLICLISLVLQFQWRYLFSKLMRMLLILMIALWFELDWSTCSQICAK 840 |
| CC | PS | Db | 781 KENIAVGLICLISLVLQFQWRYLFSKLMRMLLILMIALWFELDWSTCSQICAK 840 |
| CC | XX | Qy | 841 WTRTEAGSKKSLSSEGHHMDLPDVVITSLPGSGAEILKQLFNNSDFLYIRVTAYDI 900 |
| CC | PS | Db | 841 WTRTEAGSKKSLSSEGHHMDLPDVVITSLPGSGAEILKQLFNNSDFLYIRVTAYDI 900 |
| CC | XX | Qy | 901 PETELEIISFDVACEWKVSDIREGHFRFLRGWQSLVDTKLUQNTHLHEPGRKLQY 960 |
| CC | PS | Db | 901 PETELEIISFDVACEWKVSDIREGHFRFLRGWQSLVDTKLUQNTHLHEPGRKLQY 960 |
| CC | XX | Qy | 961 PANKDKRKPKRRESLPDQRSMKGAFDRDASYKRALRRHLVYPSARPVLSSGSNT 1020 |
| CC | PS | Db | 961 PANKDKRKPKRRESLPDQRSMKGAFDRDASYKRALRRHLVYPSARPVLSSGSNT 1020 |
| CC | XX | Qy | 1021 LKHPPQEVLGASMRLAYVDRDAYSMLYNKSPSLYSLKVNPHELAKLFKEGGKGK 1080 |
| CC | PS | Db | 1021 LKHPPQEVLGASMRLAYVDRDAYSMLYNKSPSLYSLKVNPHELAKLFKEGGKGK 1080 |
| CC | XX | Qy | 1081 CNLNSGYAFBEYEPKELSKSKSNASVSLSHMLANTAAALRINTDLPLTSYOLVKFED 1140 |
| CC | PS | Db | 1081 CNLNSGYAFBEYEPKELSKSKSNASVSLSHMLANTAAALRINTDLPLTSYOLVKFED 1140 |
| CC | XX | Qy | 1141 VHPPQKTTERIAFLGIPLSPLASLNQILPATSNLFLPYEGISPTNTNWKONLPRDE 1200 |
| CC | PS | Db | 1141 VHPPQKTTERIAFLGIPLSPLASLNQILPATSNLFLPYEGISPTNTNWKONLPRDE 1200 |
| CC | XX | Qy | 1201 IKLIENICWTLMDRGLYKPFMD 1222 |
| CC | PS | Db | 1201 IKLIENICWTLMDRGLYKPFMD 1222 |
| CC | XX | Qy | 181 GHSLTGPATADFLYNLNDNRQKYLEKWWVITEEMEYWSKRSWGLKLLHNFQATNM 240 |
| CC | PS | Db | 181 VHSLTGFATADFLYNLNDNRQKYLEKWWVITEEMEYWSKRSWGLKLLHNFQATNM 240 |

RESULT 3
ID AAB33541
ID AAB33541 standard; protein; 1212 AA.

KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis; glomerulonephritis; myasthenia gravis; Sjogren's syndrome; hypoparathyroidism; Grave's disease; anaemia; Hashimoto's disease; demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis; myocarditis; adult respiratory distress syndrome; eczema; psoriasis; asthma; hypersensitivity; rheumatic fever; tissue rejection.

Mus sp.

PN WO200106015-A1.

XX 25-JAN-2001.

XX 19-JUL-2000; 2000WO-US019741.

PR 20-JUL-1999; 99US-0144693P.

PR 13-JUN-2000; 2000US-00593828.

XX (REGE) UNIV CALIFORNIA.

PA Rosen SD, Lee JK, Hemmerich S;

XX WPI: 2001-138471/14.

DR N-PSDB; ADO2705, ADO2706.

XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for PT diagnostic and therapeutic agent screening applications.

XX PS Example 2: Fig 6B; 129pp; English.

XX The present sequence is mouse Glycosyl sulfotransferase-6 (GST-6). GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non-sulphated selectin ligand, GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in inhibiting a selectin mediated binding event. GST is useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes, glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious anaemia, demyelinating diseases, cirrhosis, ulcerative colitis, dermatis, myocarditis, regional enteritis, adult respiratory distress syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation.

XX Sequence 1207 AA;

Query Match 90.3%; Score 589.5; DB 4; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0; Mismatches 55; Indels 5; Gaps 1;

Matches 1084; Conservative 68; Mismatches 55; Indels 5; Gaps 1;

QY 11 MAMFTGHFLFLALMPAFSTFERSVSNYSEWAFTDDIDQFTKQVDFRPNOKLXKSM 70

Db 1 MAFNTEHFLFLTLMCSFTCBEVSNSYSEWAFTDDIQQLWLSQKIQDFKLNR----- 55

QY 71 LHPFLYDQEQIOMRQSKRASHLHLFRAIRSAYTMSNPPTYLPPPKHADFAAKWNEI 130

Db 56 LHPNLYFDAGDIOQTLKOKSRSTSLLHIFRAIKSAVIMSLNSPSTYLPQQPKHAEFAAKWNEI 115

QY 131 YGNNLPLPLADYCLACPEDKVAFEPVLEYMDRVGYKDVLVNAPGDEVPIGHSLTGATA 190

Db 116 YGNNLPLPLADYCLACPEDKVAFEPVLEYMDRVGYKDVLVNAFGDEVPGHSLTGATA 175

QY 191 FDFLYNLQHRRQKYLEKIVWITEEMEYYSKVRSGKCOLLNHOATNMALLGALVTG 250

Db 176 FDFLYNLQHRRQKYLEKIVWITEEMEYSKVRSGKCOLLNHOATNMALLGALVTG 235

QY 251 VDKGSKANTWKQAVYDVMEMKTMLNHITDGSJDEGVAGSYTAKSMTTOYVLAQRHENI 310

Db 236 VDKGSKANTWKQAVYDVMEMKTMLNHITDGSJDEGVAGSYTAKSMTQVFLAQRHENI 295

| | |
|----|--|
| QY | 311 NNLDNNWILKMHFWFYATLPLGFORTYGIADSNYWFYGPESQLVPLDKFILKNGAGCNWL 370 |
| Db | 296 NNFDNNWILKMHFWFYATLPLGFORTYGIADSNYWFYGPESQLVPLDKFILKNGAGCNWL 355 |
| QY | 371 AQQIRKRKPGCPMVPSSTAQRNSTLHTHEYIWKDQPLTPQQPADYGTAKIHTTPNWGVUTY 430 |
| Db | 356 AQQIRKRKPGCPMVPSSTAQRNSTLHTHEYIWKDQPLTPQQPADYGTAKIHTTPNWGVUTY 415 |
| QY | 431 GAGLPNTQNTFVFSKGKLGRGRAVTDVHOPYSWDGWSFNPGHSPNQGPBHDONSFTRAPN 490 |
| Db | 416 GGGLPNTQNTFVFSKGKLGRGRAVTDVHOPYSWDGWSFNPGHSPNQGPBHDONSFTRAPN 475 |
| QY | 491 GOVFSRALYGPPLSHNNVLYFAPSNSOCNKPWEOLGECQAOLWMTGEEVGDAGEI 550 |
| Db | 476 GOVFSRALYGPPLSHNNVLYFAPSNSOCNQPWBOLGECQAOLWMTGEEVGDAGEV 535 |
| QY | 551 ITASQHGEWVFYSGEAVSAYSSAMRLKSVYRALLNISQTLVYDHDHERQEDSPINSVSA 610 |
| Db | 536 ITAAQHQSDRMFVSGEAVSAYSSAMRLKSVYRALLNISQTLVYDHDHERQEDSPINSVSA 595 |
| QY | 611 FPHNLDIDFKYIPIKFNRYNGAMMDWDAHYKMFWDHHGNNSPMASIQEAQAEFKCR 670 |
| Db | 596 FPHNLDIDFKYIPIKFNRYNGAMMDWDAHYKMFWDHHGNNSPMASIQEAQAEFKCR 655 |
| QY | 671 WTCQFVNVTFFOMESTITRIAYFYGPYINVSSCRFIDSSNPGLOISLWNNNTHEVSVTTD 730 |
| Db | 656 WTCQFVNVTFFHMESTITRIAYFYGPYINVSSCRFIDSSSGQLOISLWNNNTHEVSVTTD 715 |
| QY | 731 YHNLKTRFENYLGGFPASVADQGQTIPFGLGTOAIKVPRVHRDRIFPFGEKPNIAVGIL 790 |
| Db | 716 YCNLKSREFSYLGGFPASVANQGQTIPFGLGTOEVINVPVRHDCKVNFPEGEKPNIAVGFL 775 |
| QY | 791 C1SLVLTQFQWFYLSRPLMRCLVCLVILVIALFIELDWWSCTCOPICAKWTRTEAESSK 850 |
| Db | 776 C1SLVLTQFQWFYLSRPLMRCLVCLVILVIALFIELDWWSCTCOPICAKWTRTEAKANE 835 |
| QY | 851 KSL-SSEGHMMLPDWVYTISLPGSGAEILKOLFNSSSFLYTRVATYDIPETEIDSF 910 |
| Db | 836 KWMISGEHVVHDLENVITSLPGSGAEILKOLFNSSSFLYTRVATYDIPETEIDSF 895 |
| QY | 911 VDACEWVYSDIRSGHFILLRGNLSQYIYDPTKHLQNLQTHLHEPNRGKLAQYFAMNKDKRK 970 |
| Db | 896 VDACEWVYSDIRSGHFILLRGNLSQYIYDPTKHLQNLQTHLHETSRSKLAQYFTTNKDKRK 955 |
| QY | 971 FKRESLPEQRSCOMKGAFDRDAEYIRALRRHLYVYPARPVLSLSSGWTLKHFQEVN 1030 |
| Db | 956 LRRESLQDQRSRIGKPPDRDAEYIRALRRHLYVYPARPVLSLSSGWTLKHFQEVN 1015 |
| QY | 1031 GASMRALYIVDRPAWVYSLMYSKPSLSSKPSLYSLKNVPEHLAKLPKIEGGKGKCNLNSGTAE 1090 |
| Db | 1016 GTSNRALLYIVDRPAWVYSLVLSKPSLSSKPSLYSLKNVPEHLAKLPKIEGGKGKCNLNSGTAE 1075 |
| QY | 1091 YEPLRKELSKSKNSAVSILSHMLANTAAALRINTDILPTSYOLVKPEDIYRPQKTER 1150 |
| Db | 1076 YESUKKELEISQNAISILSHLKVNTAAALRINTDILPTNTYLVKPFEDIYRPQKTER 1135 |
| QY | 1151 IF AFLGLPILPSLNLQLFATSPNLTFLPYEGEISPTNTNWWKQNLPDIEKJENICWT 1210 |
| Db | 1136 IF AFLGLPILPSLNLQLFATSPNLTFLPYEGEISPTNTNWWKQNLPDIEKJENICWT 1195 |
| QY | 1211 LMRLGYPKFMID 1222 |
| Db | 1196 LMRLGYPKFMID 1207 |

RESULT 5
AAE33542
ID AAE33542 standard; protein; 1207 AA.
XX AC
XX DT 16-APR-2003 (first entry)

| | | |
|--|--|---|
| DB | 301 | PSGQCNKPWGQLGEAQWKLWKTGEVGDAAGEITIASQHGMVFSGEAVSAYSAMRL 360 |
| XX | 29-JUL-1999; | 99JP-00248036. |
| PR | 27-AUG-1999; | 99JP-00300233. |
| PR | 11-JAN-2000; | 2000JP-00118776. |
| PR | 01-MAY-2000; | 2000JP-00183787. |
| PR | 09-JUN-2000; | 2000JP-00241899. |
| XX | (HELI-) HELIX RES INST. | |
| PA | Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J; Ishii S, Sugiyana T, Wakamatsu A, Nagai K, Otsuki T; | |
| XX | WPI: 2001-318749/34. | |
| DR | | |
| XX | Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. | |
| PT | PT | |
| PT | PT | |
| PT | PT | |
| PS | Claim 8; SEQ ID NO 13382; 2537pp + Sequence Listing; English. | |
| XX | The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining the full-length cDNAs easily without any specialised methods. AAH13616 to AAH13628 and AAH13633 to AAH13674 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13632 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention. | |
| SQ | Sequence 755 AA; | |
| Query Match Best Local Similarity 61.8%; Score 4037; DB 4; Length 755; | | |
| Matches 750; Conservative 2; Mismatches 0; Gaps 0; | | |
| Db | 217 MYBSKVRSWKGKOLHNHOATNMIAALLTGALVTDKGSKANTIKQAVYDVMETKMFJN 276 | |
| Db | 1 MYBSKVRSWKGKOLHNHOATNMIAALLTGALVTDKGSKANTIKQAVYDVMETKMFJN 276 | |
| Qy | 277 HIVGSLDEGVAYCSYATAKSVTQYFLAQRHENINNDNNWKLKHFMWYATLPGFORT 336 | |
| Qy | 61 HIVGSLDEGVAYCSYATAKSVTQYFLAQRHENINNDNNWKLKHFMWYATLPGFORT 120 | |
| Db | 337 VGIADSNWVFGPESQVLVDFKEFLKNGAAGNWAQIIRHRPKDGPWYSTAQWSTLH 396 | |
| Db | 121 VGIADSNWVFGPESQVLVDFKEFLKNGAAGNWAQIIRHRPKDGPWYSTAQWSTLH 180 | |
| Qy | 397 TEIYWDPLTPOPADYCTAKIHTFPWNGVVTYGALEPNQTNTFVSPSGKLGGRAYV 456 | |
| Db | 181 TEIYWDPLTPOPADYCTAKIHTFPWNGVVTYGALEPNQTNTFVSPSGKLGGRAYV 240 | |
| Qy | 457 DIVHFQPSWIDGRPSFNQGHEPDQNSPTFAPNGQVFSSEALYGPKLSHNNLVFAPS 516 | |
| Db | 241 DIVHFQPSWIDGRPSFNQGHEPDQNSPTFAPNGQVFSSEALYGPKLSHNNLVFAPS 300 | |
| Qy | PSSQCNKPWGQNLGEAQWKLWKTGEVGDAAGEITIASQGENMVSGEAVSAYSAMRL 576 | |
| CC | 577 KSVYRALLIANSQTLVVDDHTEQEDSPINSVSAPHNLIDDEFKYIPIKFMRNYGAMMD 636 | |
| CC | 361 KSVYRALLIANSQTLVVDDHTEQEDSPINSVSAPHNLIDDEFKYIPIKFMRNYGAMMD 420 | |
| Db | 637 WDAHYKMPMFDHGNSPMASIQUEARQAEEFKRWTQFVNNTFOMESTITRAYFVGPY 696 | |
| Db | 421 WDAHYKMPMFDHGNSPMASIQUEARQAEEFKRWTQFVNNTFOMESTITRAYFVGPY 480 | |
| Qy | 697 INVSSCRFDIDSNPGQISLNVNTTHVSVITVDYHLKTRNEYLGEGFGASYADQGQIT 756 | |
| Db | 481 INVSSCRFDIDSNPGQISLNVNTTHVSVITVDYHLKTRNEYLGEGFGASYADQGQIT 540 | |
| Qy | 757 RFGLGTOAIYKPVHRDRIIFPGFKENIAGVLCLSLVLTQFMERFLSPKLMRMWLLI 816 | |
| Db | 541 RFGLGTOAIYKPVHRDRIIFPGFKENIAGVLCLSLVLTQFMERFLSPKLMRMWLLI 600 | |
| Qy | 817 LVIALWFIELDWWSTCOSQPIAKWTRTEAGSKKSLISSEGHHMDLPDVVTSLPGSAAE 876 | |
| Db | 601 LVIALWFIELDWWSTCOSQPIAKWTRTEAGSKKSLISSEGHHMDLPDVVTSLPGSAAE 660 | |
| Qy | 877 ILKOLFENSDPFLYIVRPTAYIDIPSTELEIDSFVDAECKVSDIRSGHFRLLRGWLOSSL 936 | |
| Db | 661 ILKOLFENSDPFLYIVRPTAYIDIPSTELEIDSFVDAECKVSDIRSGHFRLLRGWLOSSL 720 | |
| Qy | 937 VQDTKHLQNLTHLHENRNGKLAQYFAMNKDKRK 970 | |
| Db | 721 VQDTKHLQNLTHLHENRNGKLAQYFAMNKDKRK 754 | |
| RESULT | 7 | |
| ID | ADP56673 standard; protein; 755 AA. | |
| AC | ADP56673; | |
| XX | 12-AUG-2004 (first entry) | |
| XX | Human protein which is a chondroitin 6 sulphotransferase 3 homologue. | |
| XX | cancer detection; large intestine; oesophagus; stomach; lungs; pancreas; liver; kidney; colon; human; chondroitin 6 sulphotransferase 3; CH6f3. | |
| XX | Homo sapiens. | |
| XX | JP2004147505-A. | |
| PN | NN | |
| XX | 27-MAY-2004. | |
| PD | | |
| XX | 2002JP-00312927. | |
| PP | | |
| XX | 2002JP-00312927. | |
| PR | | |
| XX | (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO. | |
| PA | (FJRE) FUJIREBIO KK. | |
| PA | (SEGK) SEIKAGAKU KOGYO CO LTD. | |
| XX | WPI: 2004-434547/41. | |
| DR | N-PSDB; ADP56672. | |
| XX | Novel nucleic acid such as DNA, useful as probe for identifying or detecting cancerous tissue of esophagus, stomach, lungs, pancreas, liver, kidney or colon, preferably large intestine tissue. | |
| XX | Example 1; SEQ ID NO 2; 24pp; Japanese. | |

colon. The current sequence is that of the human protein of the invention which is a chondroitin 6 sulphotransferase 3 (CH6T3) homologue.

Sequence 755 AA;

| | |
|----|--|
| CC | colon. The current sequence is that of the human protein of the invention which is a chondroitin 6 sulphotransferase 3 (CH6T3) homologue. |
| XX | |
| SQ | Query Match 61.8%; Score 4037; DB 8; Length 755; Best Local Similarity 99.5%; Prod. No. 0; Matches 750; Conservative 2; Mismatches 2; Indels 0; Gaps 0; |
| Qy | 217 MYEYSKVRSWGKQLHNQTAQNMIALTGALVTVGDGSKANIKWQAVDNEKTMFLN 276 1 MYEYSKVRSWGKQLHNQTAQNMIALTGALVTVGDGSKANIKWQAVDNEKTMFLS 60 |
| Db | 277 HIVGDSLDEGVAYGSYATKSIVTOYFLAQRFHINNNLDNNMILKMHFMYYATLPGFQRT 336 61 HIVGDSLDEGVAYGSYATKSIVQVFLAQRFHINNNLDNNMILKMHFMYYATLPGFQRT 120 |
| Qy | 337 VGIADSNYNWFYGPESQVLFDKFLKGNGNWLQAQITKRPKDGPMVPSTAQRNSTLH 396 121 VGIADSNYNWFYGPESQVLFDKFLKGNGNWLQAQITKRPKDGPMVPSTAQRNSTLH 180 |
| Db | 397 TEYIWIYDOLTPQPADGTAKHTFPNGWVITYGAGLPNTQNTFVFSKSKLGRAVY 456 181 TEYIWIYDOLTPQPADGTAKHTFPNGWVITYGAGLPNTQNTFVFSKSKLGRAVY 240 |
| Qy | 457 DIVHFQPYSWIDGWRSSNPFGHEHPDONSFTFAPINGQVSEALYGYPLSHLNLYVAFPS 516 241 DIVHFQPYSWIDGWRSSNPFGHEHPDONSFTFAPINGQVSEALYGYPLSHLNLYVAFPS 300 |
| Db | 517 PSSQCNKPWEQOLGECAQWLMKNTGEEVGDAEGLITASQHGENVFSGEAVSAYSSAMRL 576 301 PSSQCNKPWEQOLGECAQWLMKNTGEEVGDAEGLITASQHGENVFSGEAVSAYSSAMRL 360 |
| Qy | 577 KSVYRALLLNSQTLLVVDHIEQEDSPINSUSAFFNLIDFKEIYKPNKNRYNGAMMD 636 361 KSVYRALLLNSQTLLVVDHIEQEDSPINSUSAFFNLIDFKEIYKPNKNRYNGAMMD 420 |
| Db | 637 VWDHYRNFWFDHGNSPMAISQAEQAFKCRWTOPVNTFQEMESTITRATYFYGPY 696 421 VWDHYRNFWFDHGNSPMAISQAEQAFKCRWTOPVNTFQEMESTITRATYFYGPY 480 |
| Qy | 697 INYSSCRFIDSSNGLQSLNVNTTEHVSTVSDYHNLKTRNFYDGFQFASVADQSQIT 756 481 INYSSCRFIDSSNGLQSLNVNTTEHVSTVSDYHNLKTRNFYDGFQFASVADQSQIT 540 |
| Db | 757 RFGIGTQAIKVPKVHDRIIFPFGFKENIAVGLCISVILTFQWRFYLSRKLMWIL 816 541 RFGIGTQAIKVPKVHDRIIFPFGFKENIAVGLCISVILTFQWRFYLSRKLMWIL 600 |
| Qy | 817 LVIALWFELLDWSTCSQPIACKWTRDEAESKSLSLSEGHMDLPDVVTSLPGSAAE 876 601 LVIALWFELLDWSTCSQPIACKWTRDEAESKSLSLSEGHMDLPDVVTSLPGSAAE 660 |
| Db | 877 ILKOLFENSSDFLYTRVTAIDIPETELEIDSFDACEWVSDIIRSIGHFRGLRGNQLSL 936 661 ILKOPFENSSDFLYTRVTAIDIPETELEIDSFDACEWVSDIIRSIGHFRGLRGNQLSL 720 |
| Qy | 937 VQDTKLHLQNLHHEPNRGKLAQYFANKDKCRK 970 721 VQDTKLHLQNLHHEPNRGKLAQYFANKDKCKK 754 |
| Db | RESULTS 8 AY72641 standard; protein; 596 AA. XX DT 02-MAY-2001 (first entry) AC Human glycosyl sulfotransferase-6 (GST-6) fragment. XX DE Human; glycosyl sulfotransferase-6; GST-6; immunosuppressive; therapy; KW |

| | |
|-----------------------|--|
| KW | selectin binding inhibitor; gene therapy; inflammation; systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes; |
| KW | polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis; |
| KW | girometulonephritis; myasthenia gravis; Sjogren's syndrome; Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia; myocarditis; adult respiratory distress syndrome; ulcerative colitis; allergic rhinitis; asthma; hypersensitivity; psoriasis; tissue rejection. |
| KW | Homo sapiens. |
| OS | |
| XX | |
| PN | WO200106015-A1. |
| XX | |
| PD | 25-JAN-2001. |
| XX | |
| PF | 19-JUL-2000; 2000WO-US019741. |
| XX | |
| PR | 20-JUL-1999; 99US-0144694P. |
| PR | 13-JUN-2000; 2000US-00593828. |
| XX | |
| PA | (REGC) UNIV CALIFORNIA. |
| XX | |
| PJ | Rosen SD, Lee JK, Hemmerich S; |
| XX | |
| WPI | 2001-138471/14. |
| DR | |
| N-PSDB | AAD02702, AAD02703. |
| XX | |
| ET | New glycosyl sulfotransferases (GST)-4Alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications. |
| XX | |
| PS | Claim 3; Fig 3; 128pp; English. |
| XX | |
| ET | New glycosyl sulfotransferases (GST)-4Alpha, GST-4beta and GST-6 for diagnostic and therapeutic agent screening applications. |
| XX | |
| XX | The present sequence is human Glycosyl sulfotransferase-6 (GST-6). |
| CC | CC fragment. GST is a type 2 membrane protein useful for inhibiting a binding event between a selectin and a selectin ligand, which comprises contacting the selectin with a non-sulphated selectin ligand, GST and a small molecular agent that inhibits the sulphation activity of GST. GST is also useful in inhibiting a selectin mediated binding event. GST is useful in gene therapy to treat disorders such as acute or chronic inflammation, systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis nodosa, polymyositis, dermatomyositis, systemic sclerosis, glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's disease, Grave's disease, hypoparathyroidism, pernicious anaemia, demyelinating diseases, cirrhosis, ulcerative colitis, dermatitis, myocarditis, regional enteritis, adult respiratory distress syndrome, infantile eczema, porokeratosis, lichen planus, allergic rhinitis, bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection during transplantation. |
| SQ | Sequence 596 AA; |
| Query Match | 48.2%; Score 3147; DB 4; Length 596; |
| Best Local Similarity | 100.0%; Pred. No. 1..e-278; |
| Matches | 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| Qy | 627 MNRYNGAMMDWDAIYKMFWDHGNSPMAISQAEQAEFKRWTOFVNVTQEMESTIT 686 Db 1 MNRYNGAMMDWDAIYKMFWDHGNSPMAISQAEQAEFKRWTOFVNVTQEMESTIT 60 |
| Qy | 687 RIAYYFGPYINVSSCRFDSSNPLQLQISLVNVNTTEHVSYTIVDYLKTRNFYLGFFF 746 Db 61 RIAYYFGPYINVSSCRFDSSNPLQLQISLVNVNTTEHVSYTIVDYLKTRNFYLGFFF 120 |
| Qy | 747 ASVADOGQITREFGLGQTAIVKPVHRDRLIFFPGFKFNIVGLICISLVLTQEMRFYLS 806 Db 121 ASVADOGQITREFGLGQTAIVKPVHRDRLIFFPGFKFNIVGLICISLVLTQEMRFYLS 180 |
| Qy | 807 FRKJMRWILVILVIALWFELLDWSTCSOPIACKWTRDEAESKSLSSSEGHMDLPDVY 866 Db 181 PRKJMRWILVILVIALWFELLDWSTCSQICAKWTRDEAESKSLSSSEGHMDLPDVY 240 |
| Qy | 867 ITSLPGSGAETLKOLFFNSDFLYTRVTAIDIPETELEIDSFDACEWVSDIIRSGHF 926 |

| | | | |
|----|------|---|------|
| Db | 241 | ITSLPGSGAIBILKOLFNSDFLYIRVPTAYIDIPETELEIDSVDACWKVSDIRSGHF | 300 |
| Dy | 927 | RLLRGWLSQVQDTKLHLQNLTHHEPNRQGKLAQYFAMNNDKCRKFRRSLPEQRSMQKG | 986 |
| Db | 301 | RLLRGWLSQVQDTKLHLQNLTHHEPNRQGKLAQYFAMNNDKCRKFRRSLPEQRSMQKG | 360 |
| Dy | 987 | AFRDDEAYTRALRHLYVTPSARVLSSLSGSWTLKLHFFQEVLGASMRALYIVRDPRW | 1046 |
| Db | 361 | AFRDDEAYTRALRHLYVTPSARVLSSLSGSWTLKLHFFQEVLGASMRALYIVRDPRW | 420 |
| Dy | 1047 | IYSMLYNKSPKSLVSLKVNPKBPEHLAKLFLIEGGKGKCNLNSGYAFYEPLAKELSKSKSNAV | 1106 |
| Db | 421 | IYSMLYNKSPKSLVSLKVNPKBPEHLAKLFLIEGGKGKCNLNSGYAFYEPLAKELSKSKSNAV | 480 |
| Dy | 1107 | SLLSHWLWLNATAAALRINTDLPPTSYQLVKPFDIYHPPQTTERFAFLGIPLSPASLNO | 1166 |
| Db | 481 | SLLSHWLWLNATAAALRINTDLPPTSYQLVKPFDIYHPPQTTERFAFLGIPLSPASLNO | 540 |
| Dy | 1167 | ILPATSTNLFLYLPYEGEISPTNTNWKQNLPDSEKLIENICWTMLDRGYPKMD | 1222 |
| Db | 541 | ILPATSTNLFLYLPYEGEISPTNTNWKQNLPDSEKLIENICWTMLDRGYPKMD | 596 |

RESULT 9

| ID | AB94410 standard; protein; 480 AA. |
|-----|---|
| XX | AB94410; |
| AC | AB94410; |
| XX | 26-JUN-2001 (first entry) |
| XX | Human protein sequence SEQ ID NO:14997. |
| DE | Human; primer; detection; diagnosis; antisense therapy; gene therapy. |
| KW | XX |
| OS | Homo sapiens. |
| XX | EP1074617-A2. |
| PN | 07-FEB-2001. |
| XX | 28-JUL-2000; 2000EB-00116126. |
| PF | 29-JUL-1999; 990JP-00248036. |
| XX | 27-AUG-1999; 990JP-00100253. |
| PR | 11-JAN-2000; 2000JP-00118776. |
| XX | 02-MAY-2000; 2000JP-00183767. |
| PR | 09-JUN-2000; 2000JP-00241899. |
| XX | (HELI-) HELIX RES INST. |
| WPI | Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J; |
| PI | Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T; |
| XX | WPI; 2001-31874/9/34. |
| CC | Claim 8; SEQ ID NO 14997; 2537pp + Sequence Listing; English. |
| CC | The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: |
| CC | (a) an oligo-dT primer and an oligonucleotide complementary to the |
| CC | complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the |
| CC | oligonucleotide comprises at least 15 nucleotides; or (b) a combination |
| CC | of an oligonucleotide comprising a sequence complementary to the |
| CC | complementary strand of a polynucleotide which comprises a 5'-end |
| CC | sequence and an oligonucleotide comprising a sequence complementary to a |
| CC | polynucleotide which comprises a 3'-end sequence, where the |
| CC | sequence and an oligonucleotide comprising a sequence complementary to a |

| Sequence 480 AA: | | | | | | |
|------------------|--|--|------------|----------|--------|-----|
| | Query Match | Score | DB 4; | Length | 480; | |
| | Best Local Similarity | 39.8* | | | | |
| | Matches | 99.8* | Pred. No. | 2.4e-28; | | |
| | Conservative | 0; | Mismatches | 1; | Indels | 0; |
| | | | | | Gaps | 0; |
| Qy | 320 | MEFWPFYATLPGFORTVGIADSNNMFYGPESQLYFLDKPFLIKNGAGNWLAQQIRKHRP | | | | 379 |
| Db | 1 | MEFWPFYATLPGFORTVGIADSNNMFYGPESQLYFLDKPFLIKNGAENWLAQQIRKHRP | | | | 60 |
| Qy | 380 | KDGPMPVPSAQWSTLHTEIYIWDPQLTPOPPADYGTAKIHTFPNMGVVTYGAGLPTNTQ | | | | 439 |
| Db | 61 | KDGPMPVPSAQWSTLHTEIYIWDPQLTPOPPADYGTAKIHTFPNMGVVTYGAGLPTNTQ | | | | 120 |
| Qy | 440 | NTFVSPFSKGKLGRAYTDIVHFQPKSNIDGMRSENPGHEHPDNSTETAPNGCVFVSFAL | | | | 499 |
| Db | 121 | NTFVSPFSKGKLGRAYTDIVHFQPKSNIDGMRSENPGHEHPDNSTETAPNGCVFVSFAL | | | | 180 |
| Qy | 500 | YCPKLSHLNNYLVPAPPSSOCNKPKNGOLGECAQWLKWTGEBVGDAEGEITASQHCEM | | | | 559 |
| Db | 181 | YCPKLSHLNNYLVPAPPSSOCNKPKNGOLGECAQNLKWTGEBVGDAEGEITASQHCEM | | | | 240 |
| Qy | 560 | VFVSGAEVSAVSSAMRLKSVTLLIUNSOULVVDHIERQEDSPINSVSAFFHNLDDF | | | | 619 |
| Db | 241 | VFVSGAEVSAVSSAMRLKSVTLLIUNSOULVVDHIERQEDSPINSVSAFFHNLDDF | | | | 300 |
| Qy | 620 | KYIPKPFMNRNYGAMMDWDHYKMFWMFDHHGNSPMASIQEAEQQAABFKKRKTQFVNNTF | | | | 679 |
| Db | 301 | KYIPKPFMNRNYGAMMDWDHYKMFWMFDHHGNSPMASIQEAEQQAABFKKRKTQFVNNTF | | | | 360 |
| Qy | 680 | QMESTITRIAVYFGPYINVSSCRFDSSNPGLQISLNVNNTEHVVSIVTDYHNLKTRFN | | | | 739 |
| Db | 361 | QMESTITRIAVYFGPYINVSSCRFDSSNPGLQISLNVNNTEHVVSIVTDYHNLKTRFN | | | | 420 |
| Qy | 740 | YLFGFGPASVADQGQTTRFGJGTOAIVKPVHDRIIPPGKFNIAVGLILCISLVLITF | | | | 799 |
| Db | 421 | YLFGFGPASVADQGQTTRFGJGTOAIVKPVHDRIIPPGKFNIAVGLILCISLVLITF | | | | 480 |
| RESULT 10 | | | | | | |
| ID | AAY51120 | standard; protein; | 958 | AA. | | |
| XX | | | | | | |
| AC | AAY51120; | | | | | |
| XX | | | | | | |
| DT | 24-MAR-2000 | (first entry) | | | | |
| XX | | | | | | |
| DE | Human SART-2 protein. | | | | | |
| XX | | | | | | |
| KW | SART-2; human; tumor antigen protein; T cell; MHC class I antigen; | | | | | |
| KW | major histocompatibility complex. | | | | | |
| XX | | | | | | |
| OS | Homo sapiens. | | | | | |
| XX | | | | | | |
| PN | JP11318455-A. | | | | | |
| XX | | | | | | |
| PD | 24-NOV-1999. | | | | | |
| XX | | | | | | |
| PF | 08-MAY-1998. | | | | | |
| XX | | | | | | |

| | | | | | | |
|--|--|---------------------------------|----|---|--|-----------|
| PR | 08-MAY-1998; | 98JPB-00126398. | | | | RESULT 11 |
| XX | (ITOY/) | ITO Y. | | | | ADL83102 |
| PA | | | ID | ADL83102 standard; protein, 958 AA. | | |
| XX | | | XX | | | |
| DR | WPI: | 2000-05023/08. | AC | ADL83102; | | |
| DR | N-PSDB; | AAZ44181. | XX | | | |
| XX | | | DT | 17-JUN-2004 (first entry) | | |
| PT | Human cancer invasion antigen protein. | | XX | | | |
| XX | Claim 1; Page 11-15; 19pp; Japanese. | | DE | Human PRO49564, SEQ ID 304. | | |
| PS | | | XX | Immunosuppressive; Cytostatic; Antiarthritic; Antirheumatic; Antianemic; | | |
| XX | This invention describes a novel human tumor antigen protein (SART-2) recognized by T cells by combining with major histocompatibility complex (MHC) class I antigens. The tumor antigen protein can be used for the treatment and the diagnosis of tumor cells other than melanoma cell, particularly cancer. This sequence represents the human SART-2 protein described in the method of the invention, | | XX | Antiallergic; Muscular; Neuroprotective; Nephrotropic; Antiinflammatory; Gene Therapy; PRO; B cell related disorder; cancer; immune-mediated inflammatory disease; human. | | |
| CC | Sequence 958 AA; | | XX | Homo sapiens. | | |
| CC | Best Local Similarity 49.0%; Pred. No. 1.5e-150; | Score 1753.5; DB 3; Length 958; | OS | | | |
| Matches 341; Conservative 114; Mismatches 210; | Indels 31; Gaps 9; | PD | | | | |
| Qy | 72 HPSLYFDAGEIQAMRKRSRASHLHLFRAISAVTVMISNPTXYLPPKHDPAAKRNBY 131 | | XX | PPF 15-SEP-2003; 2003WO-US029097. | | |
| Db | 43 HPMYXFSRAVEAELQLRASL SHEHIAARTEAVHTMSSPPELPPNPDKYSARNNEIF 102 | | XX | PR 16-SEP-2002; 2002US-0411392P. | | |
| Qy | 132 GNNIPLAIXCLCPEDKVAFPEVLFYDMDVSYKDWWENARGDEVPIGHSLTGFATAF 191 | | XX | (GETH) GENENTECH INC. | | |
| Db | 103 GNNLGALAMFCVLYPENIEARDMKAQDYNERMAAQPSWLVKDAFWDEPLAHSLVGFATAY 162 | | PA | | | |
| Qy | 192 DFLYNLDNHRROCKYLEKIKWVITEEMYYSKVRSWKGKLHNOATNMIAJLTGALVTGV 251 | | XX | PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood WI; | | |
| Db | 163 DFLNYLSKQIQQKEFLEVANASCYMETSYRGWGFOLHHHQPTNCMAILTGSLVLYL-M 221 | | PI | WU TD; DR WPI; 2004-329389/30. | | |
| Qy | 252 DKG - SKANIWKQAVDVDMEKTMFLNLHIVDGLSDEGYAKGSYTAKSVTQYFLAQRHEN 309 | | XX | DR N-PSDB; ADL83101. | | |
| Db | 222 NOGYLQEAVYLWTKOYLTMEKSIAVLREVTGDSLYEGVAYGSYTTRSLFQMLFLYRHEN 281 | | XX | XX New PRO polypeptide, useful for diagnosing and treating a B cell related disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis. | | |
| Qy | 310 NNLDDNNWILKMHFMFYATLILPGFORTVGIADSNYVFGPSOLVFLDKFLIKNGAGAW 369 | | XX | XX The present invention relates to PRO proteins and their coding sequences. The PRO proteins are useful for diagnosing and treating a B cell related disorder, e.g. X-linked infantile hypogammaglobulinemia, polysaccharide antigen unresponsiveness, selective IgG subclasses, immunodeficiency with hyper IgM, transient hypogammaglobulinemia of infancy, Burkitt's lymphoma, intermediate lymphoma, follicular lymphoma, type II hypersensitivity, rheumatoid arthritis, autoimmune mediated haemolytic anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or ankylosing spondylitis. The PRO proteins are also useful for preparing a medicament for treating a condition that is responsive to the PRO protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO coding sequences are useful as hybridization probes in chromosome and gene mapping, in preparing PRO proteins, or in generating transgenic animals or knockout animals, which in turn are useful in the development and screening of therapeutically useful reagents. | | |
| Db | 282 INHGHPWLQKHFQAFMYRTILPGFORTVIAIDSNYVFGPSOLVFLDKFYMRRNGNW 341 | | XX | XX | | |
| Qy | 370 LAQOTKRHKPDKGPMVPSTAORMSTLHTEYIWDYDPOPLTPOPPADYGTAKHTTFPNNGVVT 429 | | XX | XX | | |
| Db | 342 LADQJRRNRVVEGGTPSKGQRNCTLHTEFLWFDGSLSKVPPDFGPTLHYFEDIGVVT 401 | | XX | XX | | |
| Qy | 430 YGAQLPNTQTNTFYSFKSGKLGGRAVYDVIHFPQY - SWIDGWRSSENFGHEHDQNSFTEA 488 | | XX | XX | | |
| Db | 402 YGSALPAEINRSFUSFKSGKLGGRAVYDVIHFRNQKOWIGKGRNENFGHEHDQNSFTEA 461 | | XX | XX | | |
| Qy | 489 PNGQVFSEALYGPKLSHINNLYVFAFSPS SOCNKWPWQGLGE -- CAQWLMKTGEEVGDA 546 | | XX | XX | | |
| Db | 462 PNGVPFITAEALYGPKYTFNNVLMFSPAVSKCSFSFSPWQVTDCCSSKWSKYKHDLAASC 521 | | XX | XX | | |
| Qy | 547 AGEIIATASOGEKMFVGSEAVASAYSSANRKLKSYTRALLINLNSOTLIVDHTEROEDSPN 606 | | SQ | Sequence 958 AA; | | |
| Db | 522 QGRVVAEEKGNGVVFIRGEVGAVNQPNLNQVNRNLLHQLLIVDQHLGESPLE 581 | | | Query Match 26.8%; Score 1753.5; DB 8; Length 958; | | |
| Qy | 607 SVSAAFFHNLDIDPKYIPIKFMRNRYGAMMDWDAHYKMFWDHGHNSPMSATOEAQAAE 666 | | | Best Local Similarity 49.0%; Pred. No. 1.5e-150; | | |
| Db | 582 TAASFFHNVDVPFEE--EYVDFGVHGAFTQRQDGLKXNWMDDTGSEKATFASVYPRG 638 | | | Matches 341; Conservative 114; Mismatches 210; Indels 31; Gaps 9; | | |
| Qy | 667 FKRWTOFVNWTOMESTITRIAVYFKGPYINVSSCRFDISSNPGLOISLNVNNTHEVVS 726 | | OY | 72 HPSLYFDAGEIQAMRKRSRASHLHLFRAIRSAVTMNLNPNTTYLPPPKHADFAAKWNERY 131 | | |
| Db | 639 YPNTGNTVNVNTMWHRSPPTRAYLFTGSPIDYQS----- FTWIGDSQQLDVF 686 | | Db | 43 HPMIIFSRRAEVAELQRLAASHEHIAARLBEAVHTMILSSPLELPWDPKDYSARWEIF 102 | | |
| Qy | 727 IVDYHNUKTRFNYL----GFGCFASV-ADGQI 755 | | Qy | 132 GNNLPLPIALYCLLCPEDKVAFEFVLEMYDMRVMGKDWLIVENAPGDEVIPCHSLTGFTATF 191 | | |
| Db | 687 IATSKHAYAT--YLWITGEATGOSAFAOVIADRHKI 719 | | Db | 103 GNNLGAALMCEVLYPNEIARDMADYMEAAQPSWLYRDAPWDPBLAHSLVGFATAY 162 | | |
| Qy | 192 DFLYNLDNHRROCKYLEKIKWVITEEMYYSKVRSWKGKLHNOATNMIAJLTGALVTGV 251 | | Qy | 163 DFLYNLISKQQKEFLEVIANASGTMYETSYRRGMGFQYLHNHOPTNCVALLTGSIUL-M 221 | | |

Qy 252 DKG--SKANIKWQAVVDVMEKTMELLNLHVTDGSLDEGVAYGSYTAKSUTQYVFLAQRHFN 309
 Db 222 NGVLOEAVLWTKTQVLTLIMKSLVILRETDGSLVEGYAVGSYTTSRLQFQMLVQRFN 281
 Qy 310 INNLDNNWLKRMHFWFYYATLIPGFTORTVADSNYNWFYGPESOLVLFDRKFLIKNGAGNW 369
 Db 282 INHFQHPWLQHQAFMVRTLIPGFTORTVADSNYNWFYGPESOLVLFDRKFLIKNGAGNW 341
 Qy 370 LAQQTRKHREPKDGMVPSTAQRWSLTHTEIWYDPQLTQPADYGTAKHTFPNGVYU 429
 Db 342 LADQIIRRNRVVEGPTPSKQWRMCTLHTEFLWYNGSLSKVSPPPDEGTPTLHYEDWGVT 401
 Qy 430 YGAGLPNTQNTTFVFSKGKLGRGAVYDTHFQPY-SWIDGWRSFNPFGHEHPDONSPTPA 488
 Db 402 YGSALPAETNRSFSEKSGKLGGRAIYDTHRNCKWDWKGRWNFNAGHEHPDONSPTPA 461
 Qy 489 PNGQVFVSEALYGPKLSHLNINVLYAPAPSISQCNKPWEGLGE-CAOWLKWTGBEVGDA 546
 Db 462 PNGVPIATEALYGPKYTFNNVLMFSPAVSKCSFSPWICQVTDSSKMSKRYKHDLAASC 521
 Qy 547 AGEITATASQIGEMYFVSGEVAISSSAMRKLSYVRALLIINSQTLIVYDHIERQEDSPIN 606
 Db 522 QGRVVAEEENGWVIFRGEGVAGYNPQLNKNVORNLHPOQLLIVQHUGBESPL 581
 Qy 607 SVSAPPHNLDIKYKIPKFMNRVNGAMMDVDAHYKMFWDHGNSPMAISQAEQAE 666
 Db 582 TAASPHNNTVDPF---ETVUDGVHGAFlRQDSIYKMMMDDTGYSERATPASVTRYRG 638
 Qy 667 FKCRKTQFYVNTFQMESTITRATYFVGPYINVSSCRFDISSNPGQSLQISLVNVNNTEHVVS 726
 Db 639 YPYNGTNVNTMHLRSPIRAAFLFIGPSIDVS-----FTVHGDQSQLDVF 686
 Qy 727 IVTDYHNLKTRENYL----GFGFASV-ADOGOI 755
 Db 687 IATSKHAYAT---YLTWGTATGQSAFAQVIAADRHKI 719

RESULT 12
 AAB33826 Standard; protein; 314 AA.
 XX AAB33826 ;
 AC ;
 DT 26-JUN-2001 (first entry)
 DB Human protein sequence SEQ ID NO:13638
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 20000EP-00116126.
 XX PR 29-JUL-1999; 99JJP-00248036.
 PR 27-AUG-1999; 99JJP-00300253.
 PR 11-JAN-2000; 20000JP-00118776.
 PR 02-MAY-2000; 20000JP-00183767.
 PR 09-JUN-2000; 20000JP-0021899.

PT diagnosis of the abnormality of the proteins encoded by the full-length
 cDNAs.
 PT Claim 8; SEQ ID NO 13638; 2537pp + Sequence Listing; English.
 PS XX
 CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo(dT) primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining the full-length cDNAs easily without any specialised methods. AAH3166 to AAH3628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB32446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX Sequence 314 AA;
 SQ Query Match 20-41; Score 1331; DB 4; Length 314;
 Best Local Similarity 32-41; Pred. No. 1..3e-112;
 Matches 313; Conservative 0; Mismatches 1; Indels 652; Gaps 1;
 YY 1 MPKCGAPPIMALMFTGHLLFLALLMPAFSTFEESVSNYSEAVFTDIDQFKEKQVQDF 60
 DB 1 MPKCGAPPIMALMFTGHLLFLALLMPAFSTFEESVSNYSEAVFTDIDQFKEKQVQDF 60
 YY 61 RPNDKIKCSMLHPSLFLYFAGEIQAMRKQSRSASHLHLFAIRSANTVMSNPNTYLPKKH 120
 DB 61 RPNDKIKCSMLHPSLFLYFAGEIQAMRKQSRSASHLHLFAIRSANTVMSNPNTYLPKKH 120
 YY 121 ADFAAKKNEIYGNNLPLPLAFLYCLCPEDKVAFFVLEYMDRMYGKDWLIVENAPGDEVPI 180
 DB 121 ADFAAKKNEIYGNNLPLPLAFLYCLCPEDKVAFFVLEYMDRMYGKDWLIVENAPGDEVPI 180
 YY 121 ADFAAKKNEIYGNNLPLPLAFLYCLCPEDKVAFFVLEYMDRMYGKDWLIVENAPGDEVPI 180
 DB 121 ADFAAKKNEIYGNNLPLPLAFLYCLCPEDKVAFFVLEYMDRMYGKDWLIVENAPGDEVPI 180
 YY 181 GHSLTGATAFDPFLYNNLDNHRRQKYLEKIKWVITEEMEYSKVRSGWKGQLHHNHOATNM 240
 DB 181 GHSLTGATAFDPFLYNNLDNHRRQKYLEKIKWVITEEMEYSKVRSGWKGQLHHNHOATNM 240
 YY 181 GHSLTGATAFDPFLYNNLDNHRRQKYLEKIKWVITEEMEYSKVRSGWKGQLHHNHOATNM 240
 DB 197 -----
 YY 241 ALIJTGALVTVGDIGKSKANJTWKQAVVDNEKTMFLLNHIVDGSLDEGVGSYTAKSVTQY 300
 DB 197 -----
 YY 361 ILKNGAGMILAQQIRKHRPKDGPMPVSTAQRWSTLHTEIXIWYDPQLTPOPADYGTAKIH 420
 DB 197 -----
 YY 481 DQNSFTFAPNGQEVSEALYGPKLSHLNINVLFAPSSSQCNKPWBQGBCAQWLKWTG 540
 DB 197 -----
 YY 421 TFPNGWVVTYGAQPLPNQTNTFVFSKGKLGGRAVYD1VHFQPYSWIDGWRSPNPGHBP 480
 DB 197 -----
 XX PA(HELI-) HBLIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or

| | | | |
|--------------------|---|---------------------------|--|
| OS | Homo sapiens. | AC | ADJ12217; |
| PN | WO200179254-A1. | XX | 20-MAY-2004 (first entry) |
| XX | | DT | |
| PD | 25-OCT-2001. | XX | Human secreted protein SeqID 71. |
| XX | | XX | |
| PF | 16-APR-2001; 2001WO-US008655. | KW | human; secreted; cancer; haemopoietic disease; anaemia; prostatitis; |
| XX | | KW | inguinal hernia; reproductive system disorder; systemic lupus erythematosus; |
| PR | 18-APR-2000; 2000US-00552929. | KW | multiple myeloma; musculoskeletal disease; gout; cardiovascular disease; fetal disease; |
| PR | 22-SEP-2000; 2000US-00568317. | KW | gout; arrhythmia; hypernatraemia; fetal alcohol syndrome; Down's syndrome; urinary incontinence; renal disorder; neural sensory disease; |
| PR | 24-OCT-2000; 2000US-00692881. | KW | Alzheimer's disease; meningitis; respiratory disease; emphysema; |
| PR | 01-DEC-2000; 2000US-00728628. | KW | occupational lung disease; endocrine disease; diabetes; |
| PR | 26-JAN-2001; 2001US-00770160. | KW | glomerulonephritis; digestive disease; portal hypertension; |
| PR | 13-FEB-2001; 2001US-00781666. | KW | irritable bowel syndrome; epithelial disease; scleroderma; |
| PR | 22-MAR-2001; 2001US-00816828. | KW | epidermolysis bullosa; cytotoxic; antianemic; antiarthritic; antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory; antipsoriatic; antibacterial; osteopathic; dermatological; antitoxin; immunomodulator; laryngeal; cardiotonic; nootropic; antilipemic; nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer; antidiabetic; anabolic; hypertensive; vulnerary. |
| XX | | XX | |
| (HYSE-) HYSEQ INC. | | OS | Homo sapiens. |
| PA | | XX | |
| XX | | XX | |
| PI | Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J; | PS | US200410132-A1. |
| PI | Yang Y, Zhao QA, Goodrich RW, Liu C, Drmanac RT, Ma Y, Wang Z; | XX | |
| PI | Weinman T; | XX | |
| XX | | XX | |
| DR | WPI; 2001-607699/69. | XX | PD 15-JAN-2004. |
| XX | | XX | |
| PS | Claim 20; SEQ ID NO 563; 153pp; English. | PS | 30-OCT-2001; 2001US-00984429. |
| XX | | XX | |
| PT | Novel polynucleotides and encoded polypeptides (protein factors, including, e.g., cytokines, such as lymphokines, interferons, and circulating soluble factors) useful for treating, e.g., Parkinson's, Alzheimer's, HIV and cancer. | PT | 09-OCT-1997; 97US-0061463P. |
| PT | | PR | 09-OCT-1997; 97US-0061527P. |
| PT | | PR | 09-OCT-1997; 97US-0061528P. |
| PT | | PR | 09-OCT-1997; 97US-0061532P. |
| PT | | PR | 09-OCT-1997; 97US-0061536P. |
| PT | | PR | 09-OCT-1997; 97US-0071490P. |
| PT | | PR | 08-APR-1998; 98WO-US021142. |
| PT | | PR | 08-APR-1998; 98US-0028843. |
| PT | | PR | 01-NOV-2000; 2000US-0244591P. |
| XX | | XX | |
| CC | This invention relates to a novel isolated DNA sequence and the mature proteins encoded by them. The invention may be useful in the development of compositions with antiparkinsonian, haemostatic, nootropic, neuroprotective, osteopathic, anti-HIV, protozoocidal, immunosuppressive, antiarthritic, antirheumatic, antidiabetic, antiallergic, antiinflammatory, anticoagulant or cycostatic activities. In addition, the sequences of the invention may be useful for gene therapy. The invention may be useful for the development of treatments for Parkinson's, Alzheimer's, thrombocytopaenia, osteoporosis, osteoarthritis, infections (including HIV, Leishmania, malaria, and various fungal infections), autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, and insulin dependent diabetes mellitus, allergic reactions and conditions (for example food allergies, insect allergies and allergic rhinitis), coagulation disorders including haemophilia, and cancer. Note: The amino acid sequences given in table 6 (SeqID 439-584) may have in frame stop codons or possible insertions/deletions as shown in the table. The sequences allocated Seq IDs 1-438 are not provided (even by reference) in the specification. | PA (ROSE /) ROSEN C A. | |
| CC | | PA (BREW /) BREWER L A. | |
| CC | | PA (DUAN /) DUAN R D. | |
| CC | | PA (RUBE /) RUBEN S M. | |
| CC | | PA (FLOR /) FLORENCE K A. | |
| CC | | PA (GREBE /) GREENE J M. | |
| CC | | PA (YOUNG /) YOUNG P B. | |
| CC | | PA (FERRI /) FERRIE A M. | |
| CC | | PA (TUGG /) YU G. | |
| CC | | PA (FLOR /) FLORENCE C. | |
| CC | | PA (EBNE /) EBNER R. | |
| CC | | PA (OLSE /) OLSEN H. | |
| XX | | XX | |
| PI | Rosen CA, Brewer LA, Young PB, Ferrie AM, | PI | 2004-050518/09. |
| XX | | XX | |
| DR | Duan RD, Ruben SM, Yu G, Florence C, Ebner R, Olsen H; | DR | N-PSDB; ADJ12161. |
| XX | | XX | |
| PT | New isolated nucleic acids and polypeptides, useful for diagnosing, treating, preventing or ameliorating diseases or disorders e.g. cancer, anaemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's disease. | PT | |
| PT | | PT | |
| PT | | PT | |
| PS | Claim 11; SEQ ID NO 71; 286pp; English. | PS | |
| XX | | XX | |
| CC | This invention relates to novel polynucleotides encoding human secreted proteins. Specifically, it refers to the vectors, host cells, recombinant CC and synthetic methods for producing human polynucleotides, polypeptides CC and antibodies. Furthermore, it relates to screening methods to identify CC agonists and antagonists that can be used to inhibit or enhance the CC production and function of the secreted proteins. The present invention CC | CC | |

RESULT 15
ADJ12217 standard; protein; 125 AA.
ID ADJ12217

describes these compositions as useful for diagnosing, treating or preventing disorders such as cancer, haematopoietic diseases including anaemia and multiple myeloma, reproductive system disorders including prostatitis and inguinal hernia, musculoskeletal diseases including systemic lupus erythematosus, and gout, cardiovascular disease including arrhythmia and hypertension, mixed fetal diseases including fetal alcohol syndrome and Down's syndrome, excretory diseases including urinary incontinence and renal disorders, neural or sensory disease including Alzheimer's disease and meningitis, respiratory disease including emphysema and occupational lung disease, endocrine diseases including diabetes and Glomerulonephritis, digestive diseases including portal hypertension and irritable bowel syndrome and connective tissue or epithelial diseases including scleroderma and epidermolysis bullosa. As such, there are various activities such as cytotoxic, antianemic, antiarthritic, antiinflammatory, antipsoriatic, antibacterial, osteopathic, dermatological, antigout, immunomodulator, antiarrhythmic, cardiant, nootropic, antilipemic, nephrotropic, uroprotective, neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive and vulnerary. This polypeptide is a human secreted protein of the invention. NOTE: This sequence does not appear in the printed specification but has been obtained in electronic format from the US patent office at the following web site www.seqdata.uspto.gov/sequence.html; Document ID: 20040010132.

XX Sequence 125 AA;

| | | | | |
|-----------------------|-----------------|--|-----------|-------------|
| Query Match | 8.7% | Score 568; | DB 8; | Length 125; |
| Best Local Similarity | 100.0%; | Pred. No. 2.5e-13; | | |
| Matches 107; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 811 | MRWLLILVIALWFELDDWSTCSQICAKWRTEAGSKRSLSSEGHRMIDLPDVVITSL | 870 | |
| Db | 1 | MRWLLILVIALWFELDDWSTCSQICAKWRTEAGSKRSLSSEGHRMIDLPDVVITSL | 60 | |
| Qy | 871 | PGSGAEILKQLFFNSDFLYIRVPTAYIDIPETELEIDSFYDACEWK 917 | | |
| Db | 61 | PGSGAEILKQLFFNSDFLYIRVPTAYIDIPETELEIDSFYDACEWK 107 | | |

Search completed: June 23, 2005, 08:43:30
 Job time : 135.064 secs

| | | |
|--|-----------------------|---|
| Copyright (c) 1993 - 2005 Compugen Ltd. | GenCore version 5.1.6 | Sequence 2, Appli |
| OM protein - protein search, using SW model | Sequence 1, Appli | Sequence 1, Appli |
| Run on: June 23, 2005, 08:34:49 ; Search time 46.0594 Seconds (without alignments) | Sequence 21143, A | Sequence 21143, A |
| 1980.512 Million cell updates/sec | Sequence 15903, A | Sequence 15903, A |
| Title: US-10-697-828-15 | Sequence 28, Appli | Sequence 28, Appli |
| Perfect score: 6532 | Sequence 28, Appli | Sequence 28, Appli |
| Sequence: 1 MPKGAPPWIMALMFTGHLL.....LIENTICWTLMDRGYPKFMD 1222 | Sequence 17110, A | Sequence 17110, A |
| Scoring table: BLOSUM62 | Sequence 3737, AP | Sequence 3737, AP |
| Gapext: 0.5 | Sequence 27, Appli | Sequence 27, Appli |
| Searched: | | |
| Total number of hits satisfying chosen parameters: | 513545 | |
| Minimum DB seq length: 0 | 28 | US-09-045-284A-2 |
| Maximum DB seq length: 2000000000 | 29 | US-09-100-911-1 |
| Post-processing: Minimum Match 0% | 30 | US-09-252-99A-22443 |
| Maximum Match 100% | 31 | US-09-248-796A-15003 |
| Listing first 45 summaries | 32 | US-09-234-024-28 |
| Database : Issued Patents AA:* | 33 | PCT-US94-07902-28 |
| 1: /cgm2_6/ptodata/1/iaa/5A_COMB.pep:* | 34 | US-09-248-796A-17210 |
| 2: /cgm2_6/ptodata/1/iaa/5B_COMB.pep:* | 35 | US-09-114-001C-3757 |
| 3: /cgm2_6/ptodata/1/iaa/6A_COMB.pep:* | 36 | US-09-616-519-27 |
| 4: /cgm2_6/ptodata/1/iaa/6B_COMB.pep:* | 37 | US-09-240-767-43224 |
| 5: /cgm2_6/ptodata/1/iaa/BCTUS_COMB.pep:* | 38 | US-09-725-17 |
| 6: /cgm2_6/ptodata/1/iaa/backfiles/pep:* | 39 | US-09-091-70-279-96 |
| Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution. | 40 | US-08-987-151-2 |
| SUMMARIES | 41 | US-09-248-796A-18722 |
| Result No. | Score | Query Match Length DB ID Description |
| 1 | 56.8 | 8.7 126 4 US-09-288-143-71 Sequence 71, Appli |
| 2 | 14.7 | 2.3 411 3 US-09-015-188-2 Sequence 2, Appli |
| 3 | 14.5 | 2.2 2504 4 US-08-851-567B-12 Sequence 12, Appli |
| 4 | 14.5 | 2.2 2504 4 US-09-817-514A-8 Sequence 8, Appli |
| 5 | 13.5 | 2.1 1844 4 US-08-851-567B-33 Sequence 53, Appli |
| 6 | 13.4 | 2.1 483 3 US-09-263-023-2 Sequence 2, Appli |
| 7 | 13.4 | 2.1 483 4 US-09-471-087-2 Sequence 2, Appli |
| 8 | 13.3 | 2.0 484 3 US-09-263-023-4 Sequence 4, Appli |
| 9 | 13.3 | 2.0 484 4 US-09-471-867-4 Sequence 4, Appli |
| 10 | 13.3 | 2.0 531 4 US-09-949-016-6471 Sequence 641, Appli |
| 11 | 13.3 | 2.0 608 4 US-09-949-016-9449 Sequence 9449, Appli |
| 12 | 12.7 | 1.9 1456 4 US-09-976-594-168 Sequence 168, Appli |
| 13 | 12.5 | 1.9 395 4 US-09-949-016-7011 Sequence 7011, Appli |
| 14 | 120.5 | 1.8 479 2 US-08-949-014-2 Sequence 2, Appli |
| 15 | 119.5 | 1.8 390 4 US-09-949-016-8813 Sequence 6813, Appli |
| 16 | 119.5 | 1.8 431 4 US-09-949-016-8893 Sequence 8893, Appli |
| 17 | 115.5 | 1.8 386 4 US-09-786-240-11 Sequence 11, Appli |
| 18 | 115 | 1.8 893 3 US-09-514-302-4 Sequence 4, Appli |
| 19 | 115 | 1.8 893 4 US-10-014-316-4 Sequence 4, Appli |
| 20 | 115 | 1.8 1938 3 US-09-514-302-2 Sequence 2, Appli |
| 21 | 115 | 1.8 1938 4 US-10-014-436-2 Sequence 2, Appli |
| 22 | 112 | 1.7 772 1 US-08-258-039A-2 Sequence 2, Appli |
| 23 | 112 | 1.7 772 2 US-08-900-051-2 Sequence 2, Appli |
| 24 | 112 | 1.7 772 5 PCT-US95-07391A-2 Sequence 2, Appli |
| 25 | 111.5 | 1.7 462 2 US-08-326-180B-13 Sequence 13, Appli |
| 26 | 111.5 | 1.7 462 3 US-09-226-529-13 Sequence 13, Appli |
| 27 | 110.5 | 1.7 1455 3 US-08-840-062-5 Sequence 5, Appli |
| ALIGNMENTS | | |
| RESULT 1 | | |
| US-09-288-143-71 | | |
| ; Sequence 71, Application US/09288143 | | |
| ; Patent No. 6433139 | | |
| ; GENERAL INFORMATION: | | |
| ; APPLICANT: Brewer et al. | | |
| ; TITLE OF INVENTION: 53 Human Secreted Proteins | | |
| ; FILE REFERENCE: P2018P1 | | |
| ; CURRENT APPLICATION NUMBER: US/09/288-143 | | |
| ; CURRENT FILING DATE: 1999-04-08 | | |
| ; EARLIER APPLICATION NUMBER: PCT/US98/21142 | | |
| ; EARLIER FILING DATE: 1998-10-08 | | |
| ; EARLIER APPLICATION NUMBER: 60/061,463 | | |
| ; EARLIER FILING DATE: 1997-10-09 | | |
| ; EARLIER APPLICATION NUMBER: 60/061,529 | | |
| ; EARLIER FILING DATE: 1997-10-09 | | |
| ; EARLIER APPLICATION NUMBER: 60/071,498 | | |
| ; EARLIER FILING DATE: 1997-10-09 | | |
| ; EARLIER APPLICATION NUMBER: 60/061,527 | | |
| ; EARLIER FILING DATE: 1997-10-09 | | |
| ; EARLIER APPLICATION NUMBER: 60/061,535 | | |
| ; EARLIER FILING DATE: 1997-10-09 | | |
| ; SEQ ID NO: 71 | | |
| ; LENGTH: 126 | | |
| ; TYPE: PRT | | |
| ; ORGANISM: Homo sapiens | | |
| ; FEATURE: | | |
| ; NAME/KEY: SITE | | |
| ; LOCATION: (126) | | |
| ; OTHER INFORMATION: Xaa equals stop translation | | |
| US-09-288-143-71 | | |
| Query Match | 8.7% | Score 568; DB 4; |
| Best Local Similarity | 100.0% | Pred. No. 3 8e-49; |
| Matches 107; | Conservative 0; | Mismatches 0; Indels 0; Gaps 0; |
| Qy | 811 | MRWLLILVALWFELDWSTSQPICKWTEAEGSKKSISSEGHHMDLPDVITSL 870 |
| Db | 1 | MRWLLILVALWFELDWSTSQPICARWTEAEGSKKSISSEGHHMDLPDVITSL 60 |
| Qy | 871 | PGGAAEILKQLQFFNSDPLYIYRPTAYDIPETELEIDSFDVACEWK 917 |
| Db | 61 | PGGAAEILKQLQFFNSDPLYIYRPTAYDIPETELEIDSFDVACEWK 107 |
| RESULT 2 | | |

US-09-015-188-2
 ; Sequence 2; Application US/09015188C
 ; Parent No. 6398358
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, Kevin J
 ; APPLICANT: Tabas, Ira
 ; TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
 ; TITLE OF INVENTION: 6-Sulfotransferase
 ; FILE REFERENCE: JEFF-0211
 ; CURRENT APPLICATION NUMBER: US/09/015,188C
 ; CURRENT FILING DATE: 1998-01-29
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 2
 ; LENGTH: 411
 ; TYPE: PRT
 US-09-015-188-2
 ; ORGANISM: Homo sapiens

Query Match 2.3%; Score 147; DB 3; Length 411;
 Best Local Similarity 18.2%; Pred. No. 1.9e-05;
 Matches 89; Conservative 58; Mismatches 149; Indels 194; Gaps 16;

| | | | |
|--------|---|--|-----|
| Qy 813 | WILJLVIATLWFIEL- | -----LDVNSTC-----SOPICAKWTRTEAEGSKKSLS | 855 |
| Db 5 | WKA VLLA ASIA IQTAIR TFAKS FRTCPGLAEAGIAERLC-----EESPTFAYNLSR | 59 | |

Qy 856 EGHMIDLDPDVITSLPGSAEILKQLPFNSDLYIRVPTAYIDIPATEELIDSFVDAE 915
 Db 60 KTH-----ILLIATRGSFSYFQLNQHLOVFLPEPL-----93
 Qy 916 WKVSDIRSCHFRLRGWLQSLOVDKHLHEPNRKLAQYFAMNKDKRKFRKRRE 975
 Db 94 -----YHVNQNLIPRTQG-----KSPADRRV 115
 Qy 976 SIEPQRSMOKGAFDRDAYI-----RALRRHLYYYPSARPV-----LS 1013
 Db 116 MLGASRDLIIRSLYDCDLYFLENYIKPPVNHHTDRIFRGASRVLCSRPVCDPPGPADLV 175
 Qy 1014 LSSGSWTKLHFFOEVLGA-----SMRALYIVRD 1042
 Db 176 LEEGDCVRKGCLLNLTVAEAACERSHVAIKTVRVEPVNDLRLALVEDPRNLKVQLVRD 235
 Qy 1043 PRAMIYSMLNSKPSLSSLYKDNVPEHLAKLFKEGGKGCNNSGYAEYPLRKELJSKS 1102
 Db 236 PRGLASRSETFRDTY-----RLWRWLWGTGR----KPYNLDVTQI-TTCVCFD 279
 Qy 1103 SNAVS--LILSHWLANTAAARINTDLIPTSYOLVKPEDIVHFPQKTERIAFPLGIPL- 1159
 Db 280 SNSVSTGLMRPPWLG-----KTMVLVRYEDLARNPMKCTEE-YGFGLIPLD 325
 Qy 1160 -----SPASLNQILFATSTNLFLYFEGEISPTNTNWKNLPPDEIKIENI 1207
 Db 326 SHYARWIQNNTRGDPTLGKHKYGTVRN-----SAATAEKWFRRLSYDIVAFAQNA 375
 Qy 1208 CWTLMDRGY 1217
 Db 376 CQVWLAQGY 385

RESULT 3
 US-08-851-567B-12
 ; Sequence 12; Application US/08851567B
 ; GENERAL INFORMATION:
 ; APPLICANT: Ension, Jerald C
 ; APPLICANT: Bowen, David J
 ; APPLICANT: Petell, James
 ; APPLICANT: Fatig, Raymond
 ; APPLICANT: Schoonover, Sue
 ; APPLICANT: French-Constant, Richard
 ; APPLICANT: Rocheleau, Thomas A.
 ; APPLICANT: Blackburn, Michael B.

Qy ; APPLICANT: Hey, Timothy D.
 ; APPLICANT: Merlo, Donald J.
 ; APPLICANT: Orr, Gregory L.
 ; APPLICANT: Roberts, Jean L.
 ; APPLICANT: Strickland, James A.
 ; APPLICANT: Guo, Lining
 ; APPLICANT: Ciche, Todd A.
 ; APPLICANT: Sukhapanda, Kititard
 ; TITLE OF INVENTION: Insecticidal Protein Toxins From Photorhabdus
 ; NUMBER OF SEQUENCES: 88
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dow AgroSciences Patent Department
 ; STREET: 9330 Zionsville Road
 ; CITY: Indianapolis
 ; STATE: IN
 ; COUNTRY: US
 ; ZIP: 46268
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/851,567B
 ; FILING DATE: 05-MAY-1997
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/063,615
 ; FILING DATE: 18-MAY-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/395,497
 ; FILING DATE: 28-FEB-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/007,255
 ; FILING DATE: 06-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/608,423
 ; FILING DATE: 28-AUG-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/705,484
 ; FILING DATE: 26-AUG-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seay, Nicholas J
 ; REGISTRATION NUMBER: 27386
 ; REFERENCE/DOCKET NUMBER: 960296.93804
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 608-251-5100
 ; TELEFAX: 608-251-9166
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2504 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-851-567B-12

Query Match 2.2%; Score 145; DB 4; Length 2504;
 Best Local Similarity 18.1%; Pred. No. 0.00091;
 Matches 231; Conservative 180; Mismatches 443; Indels 422; Gaps 60;

Qy 29 FSTFRESVSNYSEWA-----VFTFDIDQFPTOKYDFRNQKLKKSMLHPSLYF 78
 Db 1003 FTDWARYKRYSTVAVGSELVYPENYDPTQRIGQTQMD-ALLQSINQSUNADTV 1061

Qy 79 AGEI0AMRQKSRAHNLHFAIRSAVT-----MLSNPYTYLPPPKHA----D 122
 Db 1062 AFKTY-LTSFEQVNLKVISAYHNVNVDQGLTYFIGDQAAPTCYYNSVDSHKCENGK 1120

Qy 123 FAA---KVNENIYGNNLPLALYCLCPEDKVAEPFVLYMDRAGYKOWLVENAPGDEV 178
 Db 1121 FAANAWGENKI-----TCAVNPWKNIIRPVV-YMSRL-YLLWLEQOSKKSD- 1165

Qy 179 PIGHSLTGATADFPLYNLLDNRHRQKYLEKIV----VITEEMEYESKVRSMGKQULH 232

| | | | | |
|----|------|---|------|---|
| Db | 1166 | -----DGKRTIYQ--YNLKLAAHTR---YDGSWNTTPFTDVTKEVKNYTSSTDAESL-- | 1212 | 2002 --DLPKAPLTIHFPQ 2015 |
| Qy | 233 | NHQATNMIALLTGALVTGVKGSKANIWKAVDVMEKTMFLNLHIVDGSLEDEVAXSY 292 | | RESULT 4 |
| Db | 1213 | -----GLYCTGYG-----EDTLYMYSMQSS-----YESY 1239 | | US-09-817-51A-B |
| Qy | 293 | T---AKSVTQVFLAQRFHENINLD--NNWLKMGHFWFYATLLEPGFORTGIADSNTMWF 347 | | ; Sequence 8, Application US/09817514A |
| Db | 1240 | TDNNAPVTGUYIFADMSSDMTNQAQTNYNNSY-----EQFDTUMADPSSD--- 1286 | | ; Patent No. 6639129 |
| Qy | 348 | YGPESQLVFLDKFLRKGAGNWLAQKIRKPRKDPMVPQAQRWSTLHTIEIWYDPOU 407 | | ; GENERAL INFORMATION: |
| Db | 1287 | :---NKCVITRRVNNRYAEDY-----IPSSV---TSNSNYSWGDHSLT 1323 | | ; APPLICANT: ffrench-constant, Richard |
| Qy | 408 | PQPADYGTAKHTPWNWGTYGGLP-----TOTNTFVFS-FKSGLKGGRAY 455 | | ; APPLICANT: Bowen, David |
| Db | 1324 | :---MLYGSVPNNTFESEAIDLRLSTNMASTIHNNGTAGTRI 1363 | | ; APPLICANT: Rochleau, Thomas |
| Qy | 456 | YD1VHFQPKSWIDGWRSFNQHGHEPDONSTTPAPEGQVFTSEALYGPKLSHLNVLVFAP 515 | | ; APPLICANT: Waterfield, Nicholas |
| Db | 1364 | QCNLMKQYASLGDKTIIYDSEFD--DANRPNLVP---LF---RGFKDENSDDSCICINB 1414 | | ; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS |
| Qy | 516 | SPSSOCNKPWEGOLGBCAQWLKWTEB---EVGDAAGEITIASORGEMYEVSGRAVSAYS 571 | | ; FILE REFERENCE: 6165 |
| Db | 1415 | NPSS-E-DKWWFSSKDDNKTADYNGGTQC1DAGTSNKDFYLNQLEIEVISTGGYWSYYK 1473 | | ; CURRENT APPLICATION NUMBER: US/09/817 , 514A |
| Qy | 572 | SAMRKSVYARLLLNQSOTLWVHDIERQDSPINSVAFFHNLIDFCKYKPFMNRYN 631 | | ; CURRENT FILING DATE: 2000-03-26 |
| Db | 1474 | G-----NPINI-----NTGDSARV--KVTKAG 1496 | | ; PRIOR APPLICATION NUMBER: US 60/191806 |
| Qy | 632 | GAMMDVWDALKMFWDHHGNSPMASIQUEAQBAAEQQNFNLTDCKNLFNDQ 1542 | | ; PRIOR FILING DATE: 2000-03-24 |
| Db | 1497 | G-----DDQIPTADNSSTYVPO-----QAPSFEEEMYQFNNTLICKNLFNDQ | | ; NUMBER OF SEQ ID NOS: 8 |
| Qy | 690 | YVFYGPBYINVSSCRFDSSNPGQLQISLNNTVNTEHVSIYTVDHNL-----KTRFENYL 741 | | ; SEQ ID NO: 8 |
| Db | 1543 | HIEIDFTATAQDGREFLGAETFLIPVKRQVJGNTVFLRNL 1602 | | |
| Qy | 742 | GFGGPASVADOG-----QITRFLGTOIAVKPVRHDLIFPPFKENIAGU 789 | | |
| Db | 1603 | FAQOLVSRANGRIDAVSMTQNQIEPQBLGAGTYQVLKDYIDESIHTGNKSFALEYDI 1662 | | |
| Qy | 790 | LCISLVLVLTWRFLPSFRKLMRILVIALWPELL---DWSTSCOPICAKW--- 841 | | |
| Db | 1663 | FKENDSFVIQGELSETSQTVK---VPLSYPLTEATGKHNHLW---VRKYQKET 1711 | | |
| Qy | 842 | -----TRBAE-----GSKKSLSSEGHHMDLPDVWITSLPGSGA--- 875 | | |
| Db | 1712 | TDKLFDRDEKDPHGWFLSDDHKTFSGLOSSAQLKNDSSPEMDF-----SGNAL 1761 | | |
| Qy | 876 | -----BILKQFLFNSSD--GSKKSLSSEGHHMDLPDVWITSLPGSGA--- 918 | | |
| Db | 1762 | YFWELLYTTPMMAHRLLOQEFDANHWRVYNSPSGYI-----VKGKIAIYHNNV 1813 | | |
| Qy | 919 | SDIRSHFRLLRGWLOSIVQDTKLHQINLHBN-----RGKLAQPA----NN 964 | | |
| Db | 1814 | -----RPLEDTSNNAQQLDSTDPAQDPMHYKVATMPATDLIMAR 1838 | | |
| Qy | 965 | KDKRKFKKRRESLPEORSONGAFDRDAEYTRALR----RHLYYPSARPVLSSGS 1018 | | |
| Db | 1859 | GDAAYRQLERDTLAEAKMW-----VTOALNLLGDBPQVMULSTTWANPTLGNAAASK 1908 | | |
| Qy | 1019 | WTKLHFFQEVIGASHRVALYVDRPAWISMSLYNKSPSLXSYSLKNVPEHIAKLPKIEGSK 1078 | | |
| Db | 1909 | TTQVQR--QOVL-TQRLRNRSVKT-----LLGTPANSSTFLPOENSKL-----K 1951 | | |
| Qy | 1079 | G-----KCNLNSGYAFTBPLRKELSKSKSNAVSLLSHMLWANTAALRINTDLPT 1130 | | |
| Db | 1952 | GYWRTLAQRMNLRHNLSDQPLSUPLYAKPADPKALLSAAVASASQGGA----- 2001 | | |
| Qy | 1131 | SYQLVKFEDIVH-FPQ 1145 | | |
| Db | 1474 | IS----- | | |

Qy 632 GAMNDVWDAAHYKMEWFDDHGNSPMAISQEAEQAFKKRWTQFVNNTFQMEST--ITRIA 689
 Db 1497 G-----DDQITTAIDNSTYVHQ-----QPAPSFEIMIQNNNLTDCKNLFIDNQA 1542
 Qy 690 YVFYGPYINVSSCRFDSSNPGLOIISLNVNNTREHVSIVTVDYHNL-----KTRFNYL 741
 Db 1543 HIEIDFTATAQDRFLGATEFFIIPVTKVLTGTENIALYSENNGVQYMOIGAYTRTLNTL 1602
 Qy 742 GFGFPASVADQ-----QITRFGFLGTOAIVKPVRAHRIIPPPGFKFENIANGLI 789
 Db 1603 FAQQLVSPANRGIDAVLSMTEQIOPQLGAGTYVQLVLDKYDESTHTGNNSFAIEYVDI 1662
 Qy 790 LCISLVLITFQWRFYLSPRKLMERWILILWIFIELI----DWSSTCSQPICAKW---- 841
 Db 1663 FKENDSFVYVQELSETOTVVK-----VFLSYFIEATGNQHHLW-----VRAKYQET 1711
 Qy 842 -----TRTEAE-----GSKKSLSSEGHHMDLDPDVVITSLPGSGA---- 875
 Db 1712 TDKILFDRTPDEKOPHGWFLSDDHRTFGSLSAQALKNDSEPMDF-----SGNAL 1761
 Qy 876 -----EILKQLFNNSSD-----FLYLRVPAYIDPTELEIDSVDACEWKY 918
 Db 1762 YFWELFYXTTPMMMAHRLQEQNFDAAHNHWRYWPSPCYI-----VGDKTIAIYHWNV 1813
 Qy 919 SDIRSFGHFRLLRGWLQSLVQDTKLUHONIHLLH-----RGKLAQYPA----MN 964
 Db 1814 -----RPLEEDTSWNAQQLDSTDPAVAQDDPMHYKUTATFATDILMAR 1858
 Qy 965 KDKKRFKFKRRESLPEQRSMQGAFDRDAEYTRALR-----RHLVYTPPSAREVLSLSSGS 1018
 Db 1859 GDAAYRQLERDTLAEKHW-----YTQALNLUGDDEPOVMLSTTWANTLGNAAK 1908
 Qy 1019 WTLKLHFFQEVLGASMRLAYIVDRPAMISMLYNKSPSLSIKNTVPEHLAKLFKEGGK 1078
 Db 1909 TTQOYR-----QQLV-TQLRLNRSRVT-----LLGTRANSITL-FLPOENSKL-----K 1951
 Qy 1079 G-----KCMNSGSAFEYPEPKRLESKSKSNAVESSHLWMLANTAAALRINTDLPT 1130
 Db 1952 GYVRTLQAQMFMNRHNLSTGQPLSLPLYAKPADPKAULSAAVSASOGGA----- 2001
 Qy 1131 SYOLVKFEDIVH-FPQ 1145
 Db 2002 --DLPKAPLTHRFPQ 2015

Query Match 2.1%; Score 135; DB 4; Length 1844;
 Best Local Similarity 17.6%; Pred. No. 0.005;
 Matches 186; Conservative 150; Mismatches 365; Indels 354; Gaps 48;

Qy 29 FSTFEESVSNYSEWA-----VFTEDDIDQFKTQKVOQDFRPNQKLKQKSMLHPSLYFID 78
 Db 916 FTDWERYNKRYSFTWAGVSELVYPPENYDPQTQIGOTKMD-ALLQSINOSQLNADTVED 974
 Qy 79 AGEIQAMROKSRAJSHLHFLAIRSAVTV-----MLSNPPTYLPPPQPKHA----D 122
 Db 975 AFKTY-LTSEFQVANLKVTSAYHDNNVNDGQGLTYFIGDQDAAPGTYWSDHSKCCENGK 1033
 Qy 123 PAA-----KNEIYIGNNLPLPLAUYCLCIPEDKVAFFPELYMDRKYGYKOMLIVENAPGDEV 178
 Db 1034 FAANAWGEWKLK-----TCAVNPWKNIIREVV-YMSRL-YLWLEQSKKSD-1078
 Qy 179 PIGHSLTGATAFDLYNLNDNHRQKYLEKIW-----VITEENYEYSKVRSMCKQLH 232
 Db 1079 -----DGKTTIYQ--YNLKLAHTR--YDGSMNTPPTFTVYTKWNTYTSSTDAAESTL--1125
 Qy 233 NHQATNMIALTGTALVYDQGSKXANIKWAOAVDVMEKTMFLNHNIDVGSLDEGVAYGSY 292
 Db 1126 -----GLYCTGKG-----EDTLLVMFYSMQSS-----YSSY 1152
 Qy 293 T---AKSTQTYVFLAQRHFNINNL-----VNWLRKMHEWTFYYATLPGFORTYGIADSNNYWF 347
 Db 1153 TDNNAPVFTGLYIADMDSDNMTNAQTYNTWNS-----POFTWADPBD----1199
 Qy 348 YGPESOLVFLDKFILKNGAGNWLAQIRKGRPKDGMVPSSTAQMSTMTHTEIYWDQLT 407
 Db 1200 -----NKKVITRVNNRYAEDYE-----IPSSV----TSNSNSGHDHSLT 1236

RESULT 5
 US-08-851-567B-53
 Sequence 53, Application US/08851567B
 Patent No. 652884
 GENERAL INFORMATION:
 APPLICANT: Ensign, Jerald C
 APPLICANT: Bowen, David J.
 APPLICANT: Petell, James
 APPLICANT: Patig, Raymond
 APPLICANT: Schoonover, Sue
 APPLICANT: ffrench-Constant, Richard
 APPLICANT: Rocheleau, Thomas A.
 APPLICANT: Blackburn, Michael B.
 APPLICANT: Hey, Timothy D.
 APPLICANT: Merlo, Donald J.
 APPLICANT: Orr, Gregory L.
 APPLICANT: Roberts, Jean L.
 APPLICANT: Strickland, James A.
 APPLICANT: Guo, Lining
 APPLICANT: Ciche, Todd A.
 APPLICANT: Sukhapinda, Kitisri
 TITLE OF INVENTION: Insecticidal Protein Toxins From *Photobacterabutus*
 NUMBER OF SEQUENCES: 88
 CORRESPONDENCE ADDRESS:
 ADDRESS: Dow AgroSciences Patent Department
 STREET: 9330 Indianapolis Road
 CITY: Indianapolis

Qy 408 POPPADDYGTAKIHTFPNGVTVYGAFLPN-----TQINTNFVS-FKSGRLGGRAV 455
 Db 1237 -----MLIGGSVPVNITFESSAEDIRLSTTMALSTHNGYAGTRRI 1276
 Qy 456 YD1VHFQPYSWIDGWRSEFNPGHEHPDONSFTFAPNGQVFSEALYGPKLSHUNLVYFAP 515
 Db 1277 QCNLMKQYASLGDKPIYDSDFD--DAKRNLVYD--LP---KPGKDENSDDSICYNE 1327
 Qy 516 SPSSQCNKPWEGQLGECRQLWKWTGE---EVGDAEGEITTAQHGENWVPSGEAVSAYS 571
 Db 1328 NPSSE-DKXWYFSSSKDNTKNTADYINGTQCIDAGTSNKFQYNNQIEEVISVTGGWSSYK 1386
 Qy 572 SAMRLKSVYRALLLNSQTLVVDHIERBODSPINSVAFHNDIDFKYIPVKPMARYN 631
 Db 1387 IS-----NPINI-----NTGIDSAKV--KVTYKAG 1409
 Qy 632 GAMMDVWDAAHYKMWFDHHGNSPMAISQEAEQAFKCRWTOFVNVTOMEEST--ITRIA 689
 Db 1410 G-----DDQIFTADNSTYVPQ-----QPAPSFEENIYQFNRLTDCKNLNFDNQA 1455
 Qy 690 YVFPGPYINSSCRFDSSNPGLOQISLNTNTHEVSYIYDYNL-----KTRFNYL 741
 Db 1456 HIEIDFTATAQDGRFLGAETFLIPVTKKVYLNTENIALYSENNQVYQMGIAFRTRNL 1515
 Qy 742 GFGGPASVADQ-----QITRFGLGTOAIKVEPRHRDLIIPPGFKENIAGLI 789
 Db 1516 FAQQLVLVRANGIDAVLSMETONIQEPOLQAGTVQLVLDKYDESIHGNTNSFAIEYDI 1575
 Qy 790 LCISLVLVLTQWRPFYLSFRKLMRWTLLVLYALWTELL---DWSTCSQPICAKW--- 841
 Db 1576 FKENDSFVYQGELSETSQNVK-----VELSYPLTEATGNKNHLW-----VRAKYQKET 1624
 Qy 842 -----TRTAE-----GSKSLSSEGHFMOLPDVVVITSLFGSGA--- 875
 Db 1625 TDKILFDRTKEDPHGWFLSDDHKTFSGLSQAQLKNDSEPMDF-----SGANAL 1674
 Qy 876 -----BILKOLFENPS-----BILKOLFENPS-----DSFDACEWKY 918
 Db 1675 YFWELFYTTMMMAHRLLQEQNFDANHWFYVWPSGYI-----VGDKIAIYHNNV 1726
 Qy 919 SDIRSFGHFRLLRGWNLQSLYDCTKLHLQNTLHHEPN 953
 Db 1727 -----RPLEDTSMNAQQQJLSTDPO 1746

RESULT 6
 US-09-263-023-2
 ; Sequence 2, Application US/092631023
 ; Pat. No. 6037159
 ; GENERAL INFORMATION:
 ; APPLICANT: Uchimura, Kenji
 ; APPLICANT: Muramatsu, Hideki
 ; APPLICANT: Kadomatsu, Kenji
 ; APPLICANT: Kannagi, Reiji
 ; APPLICANT: Habuchi, Osami
 ; APPLICANT: Muramatsu, Takashi
 ; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
 ; FILE REFERENCE: TOYAN41.001AU5
 ; CURRENT APPLICATION NUMBER: US/09/471, 867
 ; CURRENT FILING DATE: 1999-12-23
 ; EARLIER APPLICATION NUMBER: US 09/263, 023
 ; EARLIER FILING DATE: 1998-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-54007
 ; EARLIER FILING DATE: 1998-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-177844
 ; EARLIER FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 483
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

US-09-471-867-2
 ; Sequence 2, Application US/094711867
 ; Pat. No. 6453289
 ; GENERAL INFORMATION:
 ; APPLICANT: Uchimura, Kenji
 ; APPLICANT: Muramatsu, Hideki
 ; APPLICANT: Kadomatsu, Kenji
 ; APPLICANT: Kannagi, Reiji
 ; APPLICANT: Habuchi, Osami
 ; APPLICANT: Muramatsu, Takashi
 ; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
 ; FILE REFERENCE: TOYAN41.001AU5
 ; CURRENT APPLICATION NUMBER: US/09/471, 867
 ; CURRENT FILING DATE: 1999-12-23
 ; EARLIER APPLICATION NUMBER: US 09/263, 023
 ; EARLIER FILING DATE: 1998-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-54007
 ; EARLIER FILING DATE: 1998-03-05
 ; EARLIER APPLICATION NUMBER: JP 10-177844
 ; EARLIER FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 483
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

Query Match Similarity 2.1%; Score 134; DB 4; Length 483;
 Best Local Similarity 26.2%; Pred. No. 0.00054;
 Matches 36; Mismatches 81; Indels 32; Gaps 9;

Qy 1034 MRALYIVDRPRAWIYSMLYNSKPSLY--SLKRNPE----HLAKLFKLEGGK-GKCNLNS 1085
 Db 278 LKVHLVRDPR-A-VASSRIRSRLGGLIRESLQVRSRDPRAHMPFLEAGHKLGAKEGM 336

Qy 1086 GYAFYEPLRKELSKSKSNAVSLSHWLANTAAALRINTDLPLPTSYQLVKFEDIVHFPQ 1145
 Db 337 GPADYHAI-----GAMEVICNSMAKTLQTAQ-PDPWLGHLYVRYBDLVGDPV 386

Qy 1146 KTERIFAFGLGIPSPA---SLNQILFATSTNLFLPYEGEISPTN---TNWQNULP 1197
 Db 387 KTLRRVYDFVGLVSPMEQFALNMTSGSSSSKPFPV----VSARNTQAANAWRALT 441

Qy 1198 RDEIKLNIENICWTMUDRUGYPK 1219
 Db 442 FQQIKQVEEFCYQPMAVGYER 463

SEQ ID NO 2
 ; LENGTH: 483
 ; TYPE: PRT
 ; ORGANISM: Mus musculus

US-09-263-023-2

RESULT 8
 US-09-263-023-4
 Sequence 4, Application US/09263023
 Patent No. 6037159
 GENERAL INFORMATION:
 APPLICANT: Uchimura, Kenji
 APPLICANT: Muramatsu, Hideki
 APPLICANT: Kadomatsu, Kenji
 APPLICANT: Kannagi, Reiji
 APPLICANT: Habuchi, Osami
 APPLICANT: Muramatsu, Takashi
 TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-O-SULFOTRANSFERASE AND
 TITLE OF INVENTION: DNA ENCODING THE SAME
 FILE REFERENCE: TOYAMA1 .001AUS
 CURRENT APPLICATION NUMBER: US/09/263 , 023
 CURRENT FILING DATE: 1999-03-05
 EARLIER APPLICATION NUMBER: JP 10-54007
 EARLIER FILING DATE: 1998-03-05
 EARLIER APPLICATION NUMBER: JP 10-177844
 EARLIER FILING DATE: 1998-06-24
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: FastSEQ for Windows Version 3.0
 SEQ ID NO 4
 LENGTH: 484
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-263-023-4

Query Match 2.0%; Score 133; DB 3; Length 484;
 Best Local Similarity 26.2%; Pred. No. 0.00069;
 Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 1034 MRALYIVRDPRAWIYNSMLYNSPKSLY--SLKNVPE----HLAKLFKIEGGK-GKCNLNS 1085
 Db 279 LKVTHLRDPR-AVASSRISRHGLIRESLQVRSRDPRAHRMPFLEAGHKLGAKEGV 337

Qy 1086 GYAFYEPLRKELSKSKSNASVSLSHMLANTAAALRINTDLLPTSYOLVKFEDIVHFPQ 1145
 Db 338 GGPADYHAL-----GMEYTCNSMAKLTQALQ-PPDWLQCHLYVRYEDIVGDPV 387

Qy 1146 KTTERIFAFGLPILSPA---SINQILPATSTNLFLPYEGEISPTN---TNWVKQNL 1197
 Db 388 KTLRRVYDFGLVSPMEQFALNMNTSGSSSSKPFV----VSARNATOQANAWRTALT 442

Qy 1198 RDEIKLIENICWTLMDRGYPK 1219
 Db 443 FQQIKQVEEFCCYQPMAVLYER 464

RESULT 10
 US-09-949-016-6471
 Sequence 6471, Application US/09949016
 Patent No. 6612319
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001307
 CURRENT APPLICATION NUMBER: US/09/949 , 016
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 6471
 LENGTH: 531
 TYPE: PRT
 ORGANISM: Human
 US-09-949-016-6471

Query Match 2.0%; Score 133; DB 4; Length 531;
 Best Local Similarity 26.2%; Pred. No. 0.00082;
 Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Qy 1034 MRALYIVRDPRAWIYNSMLYNSPKSLY--SLKNVPE----HLAKLFKIEGGK-GKCNLNS 1085
 Db 326 LKVTHLRDPR-AVASSRISRHGLIRESLQVRSRDPRAHRMPFLEAGHKLGAKEGV 384

Qy 1086 GYAFYEPLRKELSKSKSNASVSLSHMLANTAAALRINTDLLPTSYOLVKFEDIVHFPQ 1145
 Db 385 GGPADYHAL-----GMEYTCNSMAKLTQALQ-PPDWLQCHLYVRYEDIVGDPV 434

Qy 1146 KTTERIFAFGLPILSPA---SINQILPATSTNLFLPYEGEISPTN---TNWVKQNL 1197
 Db 435 KTLRRVYDFGLVSPMEQFALNMNTSGSSSSKPFV----VSARNATOQANAWRTALT 489

Qy 1198 RDEIKLIENICWTLMDRGYPK 1219
 Db 490 FQQIKQVEEFCCYQPMAVLYER 511

RESULT 11
 US-09-949-016-9449
 ; Sequence 9449, Application US/0949016
 ; Patent No. 6812319
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 9449
 ; LENGTH: 608
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-9449

Query Match 2.0%; Score 133; DB 4; Length 608;
 Best Local Similarity 26.2%; Pred. No. 0.0011; Indels 32; Gaps 9;
 Matches 53; Conservative 36; Mismatches 81;

Qy 1034 MRALYIVDRPRAWIYSMLYNKSPLSY--SLKVNPK--GKCNLNS 1085
 Db 403 LKVHLVRDRA-VASSRIRSRHGLIRESLQVRSRDPRHRMPFLEAGHKLGAKKEGV 461

Qy 1086 GYAFSEYPLRKELSKSKSNAVSLSHLWANTAAALRINTLPLPTSYQVLFEDIVHEPQ 1145
 Db 462 GGPADYHAL-----GMAVEICNSMARTQLTQDQ-PDDWLOGHLYVRYEDLVGDPV 511

Qy 1146 KTERIPAFGIPISPA---SLNQILPATSTNLFLPYEGEISPTN---TNTWKQNLP 1197
 Db 512 KTLRERVYDFGFLLYSPEMBGFALNMNTSGSSSSKRFPV----VSARNATQAANAWRTALT 566

Qy 1198 RDEIKLIENICCWTLMIDRLGPK 1219
 Db 567 FQQ1KQVEEPCYQPMAVLGER 588

RESULT 12
 US-09-976-594-168
 ; Sequence 976, Application US/09976594
 ; Patent No. 673549
 ; GENERAL INFORMATION:
 ; APPLICANT: Furness, Michael
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 ; FILE REFERENCE: PA-0041 US
 ; CURRENT APPLICATION NUMBER: US/09/976, 594
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240, 409
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 1143
 ; SOFTWARE: PERL Program
 ; SEQ ID NO: 168
 ; LENGTH: 1456
 ; TYPE: PRT
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. 6673549 1674368CD1

Query Match 1.9%; Score 127; DB 4; Length 1456;
 Best Local Similarity 18.7%; Pred. No. 0.022; Indels 218; Gaps 36;

Qy 180 IGHSLTGFATADFPLYN----LLDNHRRQCYLEKIVWITEEMYEY ---SKVRSWG--- 227
 Db 518 IGHLTSTFAEANQTCNNENAYLTTIEDYEQAFLTSFGLRPERCFYFTGLSDIOTKGTFQ 577
 Qy 228 ---KQLLJHNHQATINMALLTG -ALTGTGVDKGSKANIWQAVDMEKTMFLINHIVDG 281
 Db 578 WTEEEVRFRTHWNSDMPGRKPGCVAMTRGIAGG --LWD -VLKCDEKAFKVCKWAEG 631
 282 -----SLDEGYAGYSTAKSTQYVFLAQRFH -----NIN 311
 Db 632 VTHBPKPTTPPEPKCPEDMGASSRTSLCFLYKQHGKHEKTKWPESRDFRALGDLASIN 691
 Qy 312 NLDDN---WL-----XKH -FW -FYATTLPFGORTYGIADSNNWYFGPESQVFLD 358
 Db 692 NKBEQQTIWRLITASGSYHKLFWLGLTYGSPSEGFTWSDSPVSYENVAEYEPNNYQNYE 751
 Qy 359 -----KFLLNKGAGNWLAQIIRKHR-----PKDGMVUPSTAORWSTLH 396
 Db 752 YCGBLKGDPTMSANDINCEHLNTWIC-QIQKGTQPKPEPTPAQDNP -PVTEDGW-VIY 807
 Qy 397 TEYTWYDQQLTPQPPADYGTAKTHTEPVGWVUTYAGLIPNTONTFVSPFSKGKLUGRAY 456
 Db 808 KDQYY ---FSEKETMDNARAFKRNFGDL --VSQSESEKKFLWKVNRNDQSAV 860
 Qy 457 DIVHF---QPYSWIDG-----WRSENPNGHEHPDQNSPTFAPN -----QVF 494
 Db 861 FIGILLISLDKKFAMMDSKVDYWSWATEGPNFANEDENCUTMNSNSGFNDINGCYPNAF 920
 Qy 495 VSEALYGPKLSHUN--NVLVFAFSPSSOQNKPWEGOLGECAQWKLWTKGEB----- 542
 Db 921 ICQR --HNSSINATTMTPTMPSVSPGCCKEGNFNSNKCFKLFGMFEEPERKMQEARKA 976
 Qy 543 -VGDAGGETITASCHGENMFVSGEAVSSAMRLKSYYTALLLNSQTLVVDHIERQE 601
 Db 977 CIGFG-----GNLVSQNEKEQFLTYMKDSTSASWTGLND----- 1013
 Qy 602 DSPINSVSFAFHNLIDDFKYIPYKFMNR YNG --AMMDVWDAHYKMF -----WF 647
 Db 1014 ---VNSENTF --LWTDRGRGVHTNMKGCPGGRSSLSUYEDACVUIGGASNAGKHM 1067
 Qy 648 DHHCNS-----PMASIOBAQAEFGRWTQFVNVTFOMESTIT----- 686
 Db 1068 DDTCDDSKRGYICQTRSDPSLTNPATIQ-TDGFVVKYQKSSYSLMRQFCWHEAETYCKLH 1126
 Qy 687 --RAYVFGPYINVNSCRFIDSSNPGQIOISLAVNNTER-----VVS1VTDYH 733
 Db 11127 NSLASIL-DPSNAFAWQMETSNSRVIWIAUNSLNDQYTWTDKRWRVRYTNWAADEPK 1185

RESULT 13
 US-09-919-016-7011
 ; Sequence 7011, Application US/09949016
 ; Patent No. 6812319
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949, 016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241, 755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237, 768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231, 498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 7011
 LENGTH: 395
 TYPE: PRT
 ORGANISM: Human
 US-09-949-016-7011

Query Match Similarity 1.9%; Score 123.5; DB 4; Length 395;
 Best Local Similarity 23.3%; Pred. No. 0.0043; Mismatches 53; Indels 103; Gaps 20;
 Matches 90; Conservative 53; Mismatches 140; Indels 103; Gaps 20;

Qy 865 VVTLSPGSGAELIKOLPNNSSDFLYRPTAYIDIPETELEIDSFDACEKV-SDIRS 923
 Db 43 VLVSSWRSGSFVGQLENQHPDVYMEPA-----WHWWTTLSQ 82

Qy 924 GHFPFLRGWQOSIYDVTKU-HLQNLTHHEPNRGKLAQYFAMNKDKKKRKFRRESLPFRQS 982
 Db 83 GSATLHNAVRDLVRSVFCDMDFDAYLPWRLNLSLDF-----QWAVERSALCSPPACS 136

Qy 983 QMKCAFDRDAEYRALRHLVYVPS--ARPVLSSSSESSWTLKHFFO-EVL-----G 1031
 Db 137 ---AFPRGAISSEAVCPACQSFTRAREACRSYSHVYLFYRFTNLQVYLPLSDPA 192

Qy 1032 ASMRALTYTRDPRAWIYSMLYNSKPSLVSLSKRNVEPHIAKLFKIEGG--KGKCNLNNSCYAF 1089
 Db 193 LNDRIVHLVTRDPRAVLRSP-----EQTAKALLARDNGIVLG---TNGTWV 233

Qy 1090 EYEP-LR--KELSKSKSNASVSLSHLNTANTAAALRINTDLPPTSOLVVKFEDIVHPQK 1146
 Db 234 EADGLRVRREVER-----SHVRIAE-AATLK--PPPFGRGYLVRFDLAREPLA 282

Qy 1147 TTERIFAFGILPSP--ASLNQI-----LFTATSNLFYLPYEGEISPTN-TNV 1191
 Db 283 EIRALYAFTPSLTSPOLERWIRNNTHGSGP GARREAPTSR-----NALVNSQA 332

Qy 1192 WKQNJPRDEKLIENICWTMLDRGY 1217
 Db 333 WRLHLPFAKRRVQELCAGALQLGY 358

RESULT 14
 Sequence 2, Application US/08899514
 Patent No. 5910581

GENERAL INFORMATION:
 APPLICANT: HABUCHI, OSAMI
 ADDRESS: FUKUTA, MASAKAZU
 TITLE OF INVENTION: POLYPEPTIDE OF GLYCOSAMINOGLYCAN
 TITLE OF INVENTION: SULFOTRANSFERASE ORIGINATING FROM HUMAN AND DNA CODING
 TITLE OF INVENTION: FOR THE POLYPEPTIDE
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: KNOBBE, MARTENS, OLSON & BEAR, LLP
 CITY: NEWPORT BEACH
 STATE: CALIFORNIA
 ZIP: 92660

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/899,514
 FILING DATE:
 CLASSIFICATION: 435
 PRIORITY DATA:
 PRIORITY NUMBER: US-09-949-016-6813

PRIORITY INFORMATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: DANIEL E ALTMAN
 REGISTRATION NUMBER: 34,115
 REFERENCE/DOCKET NUMBER: TOYAM21.001AUS

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 714 760 0404
 TELEFAX: 714 760 9502
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 479
 TYPE: amino acid
 MOLECULE TYPE: protein

US-08-899-14-2

Query Match Similarity 1.8%; Score 120.5; DB 2; Length 479;
 Best Local Similarity 16.7%; Pred. No. 0.012; Mismatches 64; Indels 137; Gaps 12;
 Matches 71; Conservative 64; Miematches 152; Length 479;

Qy 843 RTEAEGSKSLSSEGHHMDLPDVITSLPGSGAEILKQLPNSSDLYEVNP-----894
 Db 118 RKEEEPERPAPAVGPRH-----VLLMATTTRGSSFVGEPFNQOGNIFYFPEPLNHBERTV 172

Qy 895 -----TAYDIDPTELEBIDSFYDACEKVSDIRSIGH--FRLRGWLOSLVQ 938
 Db 173 SPBPFGANAAASALVTRDVKLQFLCDLY--LEHFITPLPEDHLTQFMRRGSSRSRCLCE 230

Qy 939 D-----TKLHLQNTHHEPNRGKLAQYFAMNKDKKRKFRRRESLPEQRQSOMKGAFDRDA 992
 Db 231 DPVCPTPVKCVFEKTHCKNRCCGPNUVTLAAEACRRK-----EHMALKAVRIRQL 280

Qy 993 EXTRALRHLYYPPSARPVLSLSSGSWTLKLHFFQEVLGASMRYLIVDOPRAWYSMLY 1052
 Db 281 EFLQPL-----ADPRLD-----RVQLVRDPAVLAARMV 312

Qy 1053 NSKPLSYSLKNVPEHIAKLFKIEGGKGCNUNSGYAFEYEPPLRKELSKSGNAVSLSHL 1112
 Db 313 -----AAGKYCTWKWLDDEGDGLREEEVQ 339

Qy 1113 WLANTAAALRINTDL-----LPTSYQLVKFEDIVHPOTTERIFAFGLIPSPASLN 1165
 Db 340 RLRGNCESIRLSAELLRQPAWLGRGYMLVRYEDVARGPLQKARENYPFAPIPLQVED 399

Qy 1166 QI-----LFATSTNLFLYLPYEGEISPTNTNNWVONLPRDEKLIENICWTMLD 1213
 Db 400 WIQKNTQAAHDGSGYSTOKN-----SSEQFEKWRFSMPFLAQVWQAPCGPAMR 449

Qy 1214 RLGY 1217
 Db 450 LFGY 453

RESULT 15
 US-09-949-016-6813
 Sequence 6813, Application US/09949016
 Patent No. 6812339

GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016
 PRIOR FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 6813
 LENGTH: 390
 TYPE: PRT
 ORGANISM: Human
 US-09-949-016-6813

| | | | | | | | | | |
|-----------------------|-------|--------------|---|------------|------|--------|--------|------|-----|
| Query | Match | 1.8% | Score | 119.5; | DB | 4; | Length | 390; | |
| Best Local Similarity | 21.6% | Pred. | No. | 0.01-0.1 | | | | | |
| Matches | 86; | Conservative | 54; | Mismatches | 163; | Indels | 95; | Gaps | 16; |
| Qy | Dy | 849 | SKKSLSLSEGHMDLPDVVITSLPGSGAEILKQLFFNSSDFLYIRP-----TA | 896 | | | | | |
| Dy | Dy | 28 | SRGPSPAGGEDRVHVLVSSWRSGSFSLQQLISQHPDFYLMEBAWVHTTISQGSAA | 87 | | | | | |
| Dy | Dy | 897 | YIDIPETELEIDSFDVACEMKVSID- -RSGHFRLLRGNQLSVDTKLHLQNITHHEP- | 952 | | | | | |
| Dy | Dy | 88 | TLMARVDIAMSIFL- -CDNDVFDAYMPQRNLSAFFNWATRSA-----LCSPP | 134 | | | | | |
| Dy | Dy | 953 | ----NRGKLAQYFAMNKDKRKPKRRESLDEQRSQMKGAFDRAEYIRALRRLHVVYPP | 1007 | | | | | |
| Dy | Dy | 135 | ACSFAPPGTISK---QDVCKTLCTROPSSLAREACRSTSHTHUVLKEVRFPNLQVYPLL | 189 | | | | | |
| Dy | Dy | 1008 | ARPVPLSLSGSWTLKLHFQBVLGASMRLAIVDPRANTYSMLYNKSPLSLYSKVNEPEH | 1067 | | | | | |
| Dy | Dy | 190 | SDPALNL-----RIVLVRDRAVRSR-SAAGPLIARDNG-----225 | | | | | | |
| Dy | Dy | 1068 | LAKLFKIEGGKGKCNLNSGYAFEYBP--LRKELSKSKSNAVSLSHLMLANTAAIRIN | 1124 | | | | | |
| Dy | Dy | 226 | -----VLTGNGKW-----VEADPHLRLREVER-----SHVRLAE-AATLK-P | 261 | | | | | |
| Dy | Dy | 1125 | TDLIPLTSYQLVKFEDIVHFFQKTTERIAFGIPILPSASINQILPATSTNLFYLYEG-E | 1183 | | | | | |
| Dy | Dy | 262 | PPFLGRGYYLVRFEDLAREPLAEITALYAYPTGLTPOLEAWIWINTHGSGIGKPEAFH | 321 | | | | | |
| Dy | Dy | 1184 | ISPTN----TNWVKQNLPRDEIKLJENICWTLMDRIGY | 1217 | | | | | |
| Dy | Dy | 3222 | TSSRNARYNSQAWHRLPFTKLRYQEVCAQALQIGY | 359 | | | | | |
| Dy | Dy | 3222 | ----- | | | | | | |

Search completed: June 23, 2005, 08:52:18

THIS PAGE BLANK (USPTO)

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|---|
| 1 | 6532 | 100.0 | 1222 | 16 US-10-697-828-15 Sequence 15, Appl |
| 2 | 6509 | 99.6 | 1222 | 16 US-10-475-446-4 Sequence 4, Appl |
| 3 | 6465 | 99.0 | 1212 | 17 US-10-479-472A-2 Sequence 2, Appl |
| 4 | 5899.5 | 90.3 | 1207 | 16 US-10-697-828-17 Sequence 17, Appl |
| 5 | 5899.5 | 90.3 | 1207 | 17 US-10-479-472A-4 Sequence 4, Appl |
| 6 | 5477 | 48.2 | 596 | 16 US-10-697-828-9 Sequence 9, Appl |
| 7 | 568 | 8.7 | 125 | 11 US-09-984-429-71 Sequence 71, Appl |
| 8 | 568 | 8.7 | 126 | 14 US-10-150-111-71 Sequence 71, Appl |
| 9 | 147 | 2.3 | 411 | 14 US-10-021-660-128 Sequence 12, Appl |
| 10 | 147 | 2.3 | 411 | 15 US-10-211-462-97 Sequence 97, Appl |
| 11 | 147 | 2.3 | 411 | 16 US-10-408-765A-395 Sequence 395, App |

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 23, 2005, 08:49:14 ; Search time 154.482 Seconds
(without alignments)
3041.886 Million cell updates/sec

Title: US-10-697-828-15
Perfect score: 6532
Sequence: 1 MPKGAPPWIMALMFTGHLI.....LIENICWTLMDRGYPKFMD 1222

Scoring table: BLOSUM22
Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 38547976 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Published Applications AA:*

1: /cgn2_6_ptodata/1/pbpaa/us07_pubcomb.pep:*

2: /cgn2_6_ptodata/1/pubba/us01_pubaa/pct_new_pub.pep:*

3: /cgn2_6_ptodata/1/pubba/us06_pubg/us06_new_pub.pep:*

4: /cgn2_6_ptodata/1/pubba/us06_pubgomb.pep:*

5: /cgn2_6_ptodata/1/pubba/us01_pubaa/pctus_pubcomb.pep:*

6: /cgn2_6_ptodata/1/pubba/us01_pubaa/pctus_pubcomb.pep:*

7: /cgn2_6_ptodata/1/pubba/us08_pub_new_pub.pep:*

8: /cgn2_6_ptodata/1/pubba/us08_pubgomb.pep:*

9: /cgn2_6_ptodata/1/pubba/us08_pubgomb.pep:*

10: /cgn2_6_ptodata/1/pubba/us09b_pubcomb.pep:*

11: /cgn2_6_ptodata/1/pubba/us09c_pubcomb.pep:*

12: /cgn2_6_ptodata/1/pubba/us09_new_pub.pep:*

13: /cgn2_6_ptodata/1/pubba/us10_pubcomb.pep:*

14: /cgn2_6_ptodata/1/pubba/us10b_pubcomb.pep:*

15: /cgn2_6_ptodata/1/pubba/us10c_pubcomb.pep:*

16: /cgn2_6_ptodata/1/pubba/us10d_pubcomb.pep:*

17: /cgn2_6_ptodata/1/pubba/us10e_pubgomb.pep:*

18: /cgn2_6_ptodata/1/pubba/us10i_new_pub.pep:*

19: /cgn2_6_ptodata/1/pubba/us11a_pubcomb.pep:*

20: /cgn2_6_ptodata/1/pubba/us11i_new_pub.pep:*

21: /cgn2_6_ptodata/1/pubba/us12_new_pub.pep:*

22: /cgn2_6_ptodata/1/pubba/us12o_pubgomb.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|---|
| 1 | 6532 | 100.0 | 1222 | 16 US-10-697-828-15 Sequence 15, Appl |
| 2 | 6509 | 99.6 | 1222 | 16 US-10-475-446-4 Sequence 4, Appl |
| 3 | 6465 | 99.0 | 1212 | 17 US-10-479-472A-2 Sequence 2, Appl |
| 4 | 5899.5 | 90.3 | 1207 | 16 US-10-697-828-17 Sequence 17, Appl |
| 5 | 5899.5 | 90.3 | 1207 | 17 US-10-479-472A-4 Sequence 4, Appl |
| 6 | 5477 | 48.2 | 596 | 16 US-10-697-828-9 Sequence 9, Appl |
| 7 | 568 | 8.7 | 125 | 11 US-09-984-429-71 Sequence 71, Appl |
| 8 | 568 | 8.7 | 126 | 14 US-10-150-111-71 Sequence 71, Appl |
| 9 | 147 | 2.3 | 411 | 14 US-10-021-660-128 Sequence 12, Appl |
| 10 | 147 | 2.3 | 411 | 15 US-10-211-462-97 Sequence 97, Appl |
| 11 | 147 | 2.3 | 411 | 16 US-10-408-765A-395 Sequence 395, App |

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|---|
| 1 | 6532 | 100.0 | 1222 | 16 US-10-697-828-15 Sequence 15, Appl |
| 2 | 6509 | 99.6 | 1222 | 16 US-10-475-446-4 Sequence 4, Appl |
| 3 | 6465 | 99.0 | 1212 | 17 US-10-479-472A-2 Sequence 2, Appl |
| 4 | 5899.5 | 90.3 | 1207 | 16 US-10-697-828-17 Sequence 17, Appl |
| 5 | 5899.5 | 90.3 | 1207 | 17 US-10-479-472A-4 Sequence 4, Appl |
| 6 | 5477 | 48.2 | 596 | 16 US-10-697-828-9 Sequence 9, Appl |
| 7 | 568 | 8.7 | 125 | 11 US-09-984-429-71 Sequence 71, Appl |
| 8 | 568 | 8.7 | 126 | 14 US-10-150-111-71 Sequence 71, Appl |
| 9 | 147 | 2.3 | 411 | 14 US-10-021-660-128 Sequence 12, Appl |
| 10 | 147 | 2.3 | 411 | 15 US-10-211-462-97 Sequence 97, Appl |
| 11 | 147 | 2.3 | 411 | 16 US-10-408-765A-395 Sequence 395, App |

RESULT 1
US-10-697-828-15
; Sequence 15, Application US/10697828
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Steven
; ATTORNEY: Lee, Jin Kyu
; ATTORNEY: Hemmerich, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfoltransferases GST-4beta, & GST-6
; FILE REFERENCE: UCAL-138D1V
; CURRENT APPLICATION NUMBER: US/10/697, 828
; CURRENT FILING DATE: 2003-10-29
; PRIORITY APPLICATION NUMBER: 09/593, 828
; PRIORITY FILING DATE: 2000-06-13
; PRIORITY APPLICATION NUMBER: 60/144, 694
; PRIORITY FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSQL for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-697-828-15

ALIGNMENTS

RESULT 1
US-10-697-828-15
; Sequence 15, Application US/10697828
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Steven
; ATTORNEY: Lee, Jin Kyu
; ATTORNEY: Hemmerich, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfoltransferases GST-4beta, & GST-6
; FILE REFERENCE: UCAL-138D1V
; CURRENT APPLICATION NUMBER: US/10/697, 828
; CURRENT FILING DATE: 2003-10-29
; PRIORITY APPLICATION NUMBER: 09/593, 828
; PRIORITY FILING DATE: 2000-06-13
; PRIORITY APPLICATION NUMBER: 60/144, 694
; PRIORITY FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSQL for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-697-828-15

| | | | | | |
|----|---|------|----|------------------------------|------|
| Db | 61 RPNQKLKSMILHPSLYFDAGEIQAQRQKSRASHLHLFRAIRSAYTVMLSNPTTYLPPPKH | 120 | Qy | 1201 IKLIENICWTLMDRLLGYPKFMD | 1222 |
| Qy | 121 ADEAAKWNTEIGNRLPPLALYCLCPEDVKAFFEVLEIYMDRMYGYKDWLNEAPGDEVPI | 180 | Db | 1201 IKLIENICWTLMDRLLGYPKFMD | 1222 |
| Db | 121 ADFAAKWNTEIGNRLPPLALYCLCPEDVKAFFEVLEIYMDRMYGYKDWLNEAPGDEVPI | 180 | | | |
| Qy | 181 GHSLTGAFATPDFLYNLNNHRKRYLEKWWITBEMEYSKTRSGWKLHNRQATNM | 240 | | RESULT 2 | |
| Db | 181 GHSLTGAFATPDFLYNLNNHRKRYLEKWWITBEMEYSKTRSGWKLHNRQATNM | 240 | | US-10-475-446-4 | |
| | Sequence 4; Application US/10475446 | | | | |
| | Publication No. US20040198651A1 | | | | |
| Qy | 241 ALLTGALVTCVDGSKSANTIKWAQAVDVMETMFLINHIVDGLSDEGVAYGSTAKSVTOY | 300 | | | |
| Db | 241 ALLTGALVTCVDGSKSANTIKWAQAVDVMETMFLINHIVDGLSDEGVAYGSTAKSVTOY | 300 | | | |
| Qy | 301 VELAQPHENINLNDNNLKHMFHWYFYATLPGFORTGIAIDSNYNWFYGPESOLVEDKF | 360 | | | |
| Db | 301 VELAQPHENINLNDNNLKHMFHWYFYATLPGFORTGIAIDSNYNWFYGPESOLVEDKF | 360 | | | |
| Qy | 361 ILRGAGNWLAQQRKHRPKDGPMPVSTAQRWSLTHTEVWDPQLTQPADYGTAKH | 420 | | | |
| Db | 361 ILRGAGNWLAQQRKHRPKDGPMPVSTAQRWSLTHTEVWDPQLTQPADYGTAKH | 420 | | | |
| Qy | 421 TFPNGWVITYGAGLPNTOTTFVFSKSKLGGRAVYDVFHFOPYSWIDGWRSPNPGHEHP | 480 | | | |
| Db | 421 TFPNGWVITYGAGLPNTOTTFVFSKSKLGGRAVYDVFHFOPYSWIDGWRSPNPGHEHP | 480 | | | |
| Qy | 481 DQNSTEAPNGQVSEALYGPKLSHLNNDLVAPSPSSQCNKPWEGLGECAQWLKTG | 540 | | | |
| Db | 481 DQNSTEAPNGQVSEALYGPKLSHLNNDLVAPSPSSQCNKPWEGLGECAQWLKTG | 540 | | | |
| Qy | 541 BEVGDAGETITAQOHGENFVSGEAVASYSSANRLKSVYRALLLNSQLLVVDHIERQ | 600 | | | |
| Db | 541 BEVGDAGETITAQOHGENFVSGEAVASYSSANRLKSVYRALLLNSQLLVVDHIERQ | 600 | | | |
| Qy | 601 EDSPINSVSAFFNLDIDFKYIPKFKMFRYNGAMMDVWDHYMFWDHGHNPMASIOE | 660 | | | |
| Db | 601 EDSPINSVSAFFNLDIDFKYIPKFKMFRYNGAMMDVWDHYMFWDHGHNPMASIOE | 660 | | | |
| Qy | 661 AEQAAEFKPKWTQVNNTQOMESEITRILAYFVGPYINVSSCRFDSSNPQLQISLYNN | 720 | | | |
| Db | 661 AEQAAEFKPKWTQVNNTQOMESEITRILAYFVGPYINVSSCRFDSSNPQLQISLYNN | 720 | | | |
| Qy | 721 TEHVSIUTYDHNKTRFNYLGFGFASYADQGQITRFLGTQAIKVYHRDIIFFPGF | 780 | | | |
| Db | 721 TEHVSIUTYDHNKTRFNYLGFGFASYADQGQITRFLGTQAIKVYHRDIIFFPGF | 780 | | | |
| Qy | 781 KFNIAVGLLICISLVLTFQWRFLYSFKLMRWILYLTVLWFELDDWSTCSQPCAK | 840 | | | |
| Db | 781 KFNIAVGLLICISLVLTFQWRFLYSFKLMRWILYLTVLWFELDDWSTCSQPCAK | 840 | | | |
| Qy | 841 WTRTEAEGSKKSLSSEGHMIDLPDVITTSLGPGSAEILKQLFNSSDFLYIRYPTAYDI | 900 | | | |
| Db | 841 WTRTEAEGSKKSLSSEGHMIDLPDVITTSLGPGSAEILKQLFNSSDFLYIRYPTAYDI | 900 | | | |
| Qy | 901 PETELEIDSFDASEKWSIDRSCHFRLGWTQSLVOPTKLHQNLHNPGRKLAQ | 960 | | | |
| Db | 901 PETELEIDSFDASEKWSIDRSCHFRLGWTQSLVOPTKLHQNLHNPGRKLAQ | 960 | | | |
| Qy | 961 FAMRKDGKRFKRRESLPEQSONKGAFRDAEYTRALRRHLYYPSARPVLSSGGWT | 1020 | | | |
| Db | 961 FAMRKDGKRFKRRESLPEQSONKGAFRDAEYTRALRRHLYYPSARPVLSSGGWT | 1020 | | | |
| Qy | 1021 LKHFFOEVLGASMBALYTVDRPAWYTSKNSLKNVPEHLAKLPKIEGGRKKG | 1080 | | | |
| Db | 1021 LKHFFOEVLGASMBALYTVDRPAWYTSKNSLKNVPEHLAKLPKIEGGRKKG | 1080 | | | |
| Qy | 1081 CNLNSGYAFFFEPARKELSKSKSNASVSLSHLNLTANTAAARLINTLPLTSTOLVKFEDI | 1140 | | | |
| Db | 1081 CNLNSGYAFFFEPARKELSKSKSNASVSLSHLNLTANTAAARLINTLPLTSTOLVKFEDI | 1140 | | | |
| Qy | 1141 VHFPOKTERIFAFGIPUSASLNOILPATSTNLFLYEGEISPTNTNWKQNLPDRE | 1200 | | | |
| Db | 1141 VHFPOKTERIFAFGIPUSASLNOILPATSTNLFLYEGEISPTNTNWKQNLPDRE | 1200 | | | |

| | | | | | | | |
|----|------|--|------|----|--|------------------------|------|
| Db | 61 | RPNQKLLKSSMLHPSLYFPAGEIOMRQKSASHLHLFAIRSAYTVMISNPTYLPPKH | 120 | Qy | 1201 | IKLIENTICWTLMDRGYPKFMD | 1222 |
| Qy | 121 | ADFAAKWNEIYGNLNPLPLAICLCPEDKVAFFEVLEYMDRMVGYKDVLIVENAPGDEVP | 180 | Db | 1201 | IKLIENTICWTLMDRGYPKFMD | 1222 |
| Db | 121 | ADFAAKWNEIYGNLNPLPLAICLCPEDKVAFFEVLEYMDRMVGYKDVLIVENAPGDEVP | 180 | | | | |
| | | RESULT 3 | | | | | |
| Qy | 181 | GHSLTGPFATFDLYNLNDNRHQKYLEKIVITTEMEYESKVRSGWQLLHNHQATNM | 240 | Db | US-10-479-472A-2 | | |
| | | | | | Sequence 2, Application US/10479472A | | |
| Db | 181 | VHSLTGPFATFDLYNLNDNRHQKYLEKIVITTEMEYESKVRSGWQLLHNHQATNM | 240 | | Publication No. US20050118581A1 | | |
| Qy | 241 | ALLTGALVTVGDKSKANTIKWQAVDYMVKMELLNHIVGSLDBGVAVSYTAKSVTQY | 300 | | GENERAL INFORMATION: | | |
| Db | 241 | ALLTGALVTVGDKSKANTIKWQAVDYMVKMELLNHIVGSLDBGVAVSYTAKSVTQY | 300 | | APPLICANT: VAN BROECKHOVEN, CHRISTINE | | |
| Qy | 301 | VFLAQRFHNINNLDDNNLQMRQKMFHYATALLPGFORTVGTADSNYNWFYGPESOLVFLDKF | 360 | | TITLE OF INVENTION: NOVEL BRAIN EXPRESSED GENE AND PROTEIN ASSOCIATED WITH | | |
| Db | 301 | VFLAQRFHNINNLDDNNLQMRQKMFHYATALLPGFORTVGTADSNYNWFYGPESOLVFLDKF | 360 | | FILE REFERENCE: JAB-1711 | | |
| Qy | 361 | ILKNGAGNWLQAQIRKRPKGMPVSTPAQRWSLHTEIVWDQLTOPADYGTAKH | 420 | | CURRENT APPLICATION NUMBER: US/10/479,472A | | |
| Db | 361 | ILKNGAGNWLQAQIRKRPKGMPVSTPAQRWSLHTEIVWDQLTOPADYGTAKH | 420 | | CURRENT FILING DATE: 2003-12-01 | | |
| Qy | 421 | TFPNWGMVTVYAGLPLNTOTNTFVFSFKSRLGGRAVYDTHFQPSWIDRSFNPGHEHP | 480 | | PRIOR APPLICATION NUMBER: PCT/EP02/06316 | | |
| Db | 421 | TFPNWGMVTVYAGLPLNTOTNTFVFSFKSRLGGRAVYDTHFQPSWIDRSFNPGHEHP | 480 | | PRIOR FILING DATE: 2002-06-06 | | |
| | | | | | PRIOR APPLICATION NUMBER: EP 01202214.1 | | |
| Qy | 481 | DONSFTFAPNGQVFVSEALYGPKLSHLANNLVAFAPSPSSQCNKPWEQOLGECAQWLKWNG | 540 | | PRIOR FILING DATE: 2001-06-11 | | |
| Db | 481 | DONSFTFAPNGQVFVSEALYGPKLSHLANNLVAFAPSPSSQCNKPWEQOLGECAQWLKWNG | 540 | | NUMBER OF SEQ ID NOS: 12 | | |
| Qy | 541 | EEVGDAAEGBTITASQHGENMFVSEAEVASYSSAMRLKSVTRALLLNSQTLVVVDHIEQ | 600 | | SOFTWARE: PatentIn Ver. 3.2 | | |
| Db | 541 | EEVGDAAEGBTITASQHGENMFVSEAEVASYSSAMRLKSVTRALLLNSQTLVVVDHIEQ | 600 | | SEQ ID NO 2 | | |
| Qy | 601 | EDSPINSVSAAFFHNLIDDFKTYPKEMNRYNGAMMDVDAHYKMWFDHGNSPMASTOE | 660 | | LENGTH: 1212; | | |
| Db | 601 | EDSPINSVSAAFFHNLIDDFKTYPKEMNRYNGAMMDVDAHYKMWFDHGNSPMASTOE | 660 | | Query Match 99.0%; Score 6465; DB 17; | | |
| Qy | 661 | AQAAAEFKRKWTQVNVTQMEKTMFLHHVGDLSDEGVAGSYTAKSYTOVFLAQHRHNF | 720 | | Best Local Similarity 99.9%; Pred. No. 0; | | |
| Db | 661 | AQAAAEFKRKWTQVNVTQMEKTMFLHHVGDLSDEGVAGSYTAKSYTOVFLAQHRHNF | 720 | | Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0 | | |
| Qy | 721 | TEHVSVITVDHNLKTRFNYLGFGCFASVADQGQITRFGIGTQAIKVPRHDRIIFPFGF | 780 | | | | |
| Db | 721 | TEHVSVITVDHNLKTRFNYLGFGCFASVADQGQITRFGIGTQAIKVPRHDRIIFPFGF | 780 | | | | |
| Qy | 781 | KENIAVGLLICSLVILTFQWREFLSFRLKMRNLLILVIALMFLDLYWISTCOPICAK | 840 | | | | |
| Db | 781 | KENIAVGLLICSLVILTFQWRFYLSFRKLMRWLLVIALMFLDLYWISTCOPICAK | 840 | | | | |
| Qy | 841 | WTRTEAESKSKSLLSEGHMMDLPWVTSUPLPGSAAEILKOLPFNSDFLYIRVPTAYD | 900 | | | | |
| Db | 841 | WTRTEAESKSKSLLSEGHMMDLPWVTSUPLPGSAAEILKOLPFNSDFLYIRVPTAYD | 900 | | | | |
| Qy | 901 | PETELEIDSFDVACEWKVSDIRSFSHERLLRGWLOSLOQDTKLQNIHLBPNRGLQAY | 960 | | | | |
| Db | 901 | PETELEIDSFDVACEWKVSDIRSFSHERLLRGWLOSLOQDTKLQNIHLBPNRGLQAY | 960 | | | | |
| Qy | 961 | FAMNKDKKRKKERRESLPEORSOMKGAFDRAEYTRALRHRHLYPPSAREPVLSSLSSGSMT | 1020 | | | | |
| Db | 961 | FAMNKDKKRKKERRESLPEORSOMKGAFDRAEYTRALRHRHLYPPSAREPVLSSLSSGSMT | 1020 | | | | |
| Qy | 1021 | LKLHFFQEVLGASMRLYTRDPRAWIYSMLYNSKPSLSSLYKNPYEHAKLFKIEGGK | 1080 | | | | |
| Db | 1021 | LKLHFFQEVLGASMRLYTRDPRAWIYSMLYNSKPSLSSLYKNPYEHAKLFKIEGGK | 1080 | | | | |
| Qy | 1081 | CNLSNGYAFPEPLRKELSKSKNSAVSLLSHMLANTAAIRINTLLPLPSYQLYTKFED | 1140 | | | | |
| Db | 1081 | CNLSNGYAFPEPLRKELSKSKNSAVSLLSHMLANTAAIRINTLLPLPSYQLYTKFED | 1140 | | | | |
| Qy | 1141 | VHFPOKTERIIFAFGIPIPSASINQILEPATSTNLFLYVEGEIISPNTNTWVKONLPRDE | 1200 | | | | |
| Db | 1141 | VHFPOKTERIIFAFGIPIPSASINQILEPATSTNLFLYVEGEIISPNTNTWVKONLPRDE | 1200 | | | | |

Qy 611 FHNLDIDEFKYIPKFMNRYNGAMMDYDAHYKMFWDHGNSPMASTQEAEQAEFFKR 670
 Db 601 FHNLDIDEFKYIPKFMNRYNGAMMDWDAHYKMFWDHGNSPMASTQEAEQAEFFKR 660
 Qy 671 WTQFVNNTFQESTITRIAVYFGPYINVSSCRIDSSNPGLQISLNVNNTTEHVSVTVD 730
 Db 661 WTQFVNNTFQEMEPITRIAVYFGPYINVSSCRIDSSNPGLQISLNVNNTTEHVSVTVD 720
 Qy 731 YHNLKTRENYLGGCFASAYDQOITRGLGTOAIVPVRHDRIIPPGFKENIAGVL 790
 Db 721 YHNLKTRENYLGGCFASAYDQOITRGLGTOAIVPVRHDRIIPPGFKENIAGVL 780
 Qy 791 CISLVLTFQWRFLSPRKLMRWTLLVLTALMFIELDWSTCSQICAKNTRTEAGSK 850
 Db 781 CISLVLTFQWRFLSPRKLMRWTLLVLTALMFIELDWSTCSQPCAKNTRTEAGSK 840
 Qy 851 KSLSEGHMMMDLPDVVITSLPGSGAEIILKQLFFNSDFLYIRVPTAYIDIPETELEDSF 910
 Db 841 KSLSEGHMMMDLPDVVITSLPGSGAEIILKQLFFNSDFLYIRVPTAYIDIPETELEDSF 900
 Qy 911 VDACEWKVSDIRSCHFRLLRGWLSQVDTCKLHQNLHPENRGKLAQYFAMNKDKKRK 970
 Db 901 VDACEWKVSDIRSCHFRLLRGWLSQVDTCKLHQNLHPENRGKLAQYFAMNKDKKRK 960
 Qy 971 FKRRBLSPFQRSONKGAFDRDASYTRALRHLYVPSARPLSLSGSWTLKLHFFQEVL 1030
 Db 961 FKRESLPFQRSONKGAFDRDASYTRALRHLYVPSARPLSLSGSWTLKLHFFQEVL 1020
 Qy 1031 GASMRLATVDRPAWYSMLYNSKPSLSPSLXKNVPEHLAKLPKIEGGKGKCNLNSGYAFE 1090
 Db 1021 GASMRLATVDRPAWYSMLYNSKPSLSPSLXKNVPEHLAKLPKIEGGKGKCNLNSGYAFE 1080
 Qy 1091 YEPURKELSKSKNAVSLSHMLANTAAARLINTDLPLTSQVLKPEDITHFPQKTER 1150
 Db 1081 YEPURKELSKSKNAVSLSHMLANTAAARLINTDLPLTSQVLKPEDITHFPQKTER 1140
 Qy 1151 IFAFLGIPIUSPASINQILPATSTNLTYLPEEGISPTNTWKKQNLPDEIKLIENTCWT 1210
 Db 1141 IFAFLGIPIUSPASINQILPATSTNLTYLPEEGISPTNTWKKQNLPDEIKLIENTCWT 1200
 Qy 1211 LMDRLGPKFMD 1222
 Db 1201 LMDRLGPKFMD 1212

RESULT 4
 US-10-697-828-17
 Sequence 17, Application US/10697828
 Publication No. US20040185546A1
 GENERAL INFORMATION:
 ; APPLICANT: Rosien, Steven
 ; APPLICANT: Hemmerich, Stefan
 ; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, & GST-6
 ; FILE REFERENCE: UCAL-13DBV
 ; CURRENT APPLICATION NUMBER: US/10/697 828
 ; CURRENT FILING DATE: 2003-10-29
 ; PRIOR APPLICATION NUMBER: 09/593, 928
 ; PRIOR FILING DATE: 2000-06-13
 ; PRIOR APPLICATION NUMBER: 60/144, 694
 ; PRIOR FILING DATE: 1999-07-20
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO: 17
 ; LENGTH: 1207
 ; TYPE: PRT
 ; ORGANISM: mouse
 US-10-697-828-17

Query Match 90.3%; Score 5899.5; DB 16; Length 1207;
 Best Local Similarity 89.4%; Prod. No. 0;
 Matches 1084; Conservative 68; Mismatches 55; Indels 5; Gaps 1;

RESULT 5
US-10-697-828-15

Sequence 4, Application US/10479472A-4
Publication No. US2005/0118581A1
GENERAL INFORMATION:
APPLICANT: DEL-FAVERO, JURGEN PETER LODE
APPLICANT: VAN BROECHOVEN, CHRISTINE LODE
TITLE OF INVENTION: NOVEL BRAIN EXPRESSED GENE AND PROTEIN ASSOCIATED WITH
FILE REFERENCE: JAB-1711
CURRENT APPLICATION NUMBER: US/10/479,472A
PRIORITY FILING DATE: 2003-12-01
PRIORITY APPLICATION NUMBER: PCT/EP02/06316
PRIORITY FILING DATE: 2002-06-06
PRIORITY FILING DATE: 2001-06-11
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 3.2
SEQ ID NO 4
LENGTH: 1207
TYPE: PRT
ORGANISM: Mus sp.
US-10-697-828-15

Query Match 90.3%; Score 5899.5; DB 17; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0; Mismatches 55; Indels 5; Gaps 1;
Matches 1084; Conservative 68;

Qy 11 MAMFTGHFLFLALLMFASTFERSVSYSEWAFTDDIDQFKTQKVODFRPNOKLKKSM 70
Db 1 MAPMFTEHLLFLTLNMCSTCEESVSNYSEWAFTDDIQQLWLSQKIQQFKLNRR----- 55

Qy 71 LHPSLYFDAGEIQAMRKRSBASHLHLFRARSATTMVSNTPTYLPPPKHADPAKNEI 130
Db 56 LHPNLYFDADCIOLKOKSRTHSHIIFRAKSATTMVSNTPLYPPEAKNEI 115

Qy 131 YGNNLPPLAYCLCPEDKVAFEVLYEMDPMGVKDWLVEAGDEVIGHSLTGATA 190
Db 116 YGNNLPPLAYCLCPEDKVAFEVLYEMDPMGVKDWLVEAGDEVIGHSLTGATA 175

Qy 191 FDFLYNLNDHRROKYLEKWIWTEEMEYSKVRSGKOLLHNHOATNMALITGALYTG 250
Db 176 FDFLYNLNDHRROKYLEKWIWTEEMEYSKVRSGKOLLHNHOATNMALITGALYTG 235

Qy 251 VDKGSKANTIKQAYDVMEKTMELNHIVGSDLEGVATGSYTAKSVTQYFLAQRFHENI 310
Db 236 VDKGSKANTIKWQVYVDMEKTMFLKHIVGSDLEGVATGSYTSKSVTOYFLAQRFHENI 295

Qy 311 NNLDNNWLKHFWFVYATLPGQFORTVGZADSNNWFGESQVFLDKFILKGAGNWL 370
Db 296 NNFDNNWLKHFWFVYATLPGQFORTVGZADSNNWFGESQVFLDKFILKGAGNWL 355

Qy 371 AQQIRKHRPKDGPMPVSTAGRWTSLHTEYIWDPOLTPPADYGTAKITHFPNWGVVTY 430
Db 356 AQQIRKHRPKDGPMPVSTAGRWTSLHTEYIWDPOLTPQPVDFGTAKHTFPNWGVVTY 415

Qy 431 GAGLIENTOTTFVSKSGKIGGRAYVYDIVFQPSWIDGRTRSFNPGEHEPDQNSFTFAPN 490
Db 416 GGGLPNTQMTFVSKSGKIGGRAYVYDIVFQPSWIDGRTRSFNPGEHEPDQNSFTFAPN 475

Qy 491 GOVFVSEALYGPKLISHIINNVLYFAPSQQCNPWPWEGOLGECAQWLKWTFGEYGDAGEI 550
Db 476 GOVFVSEALYGPKLISHIINNVLYFAPSQQCNPWPWEGOLGECAQWLKWTFGEYGDAGEI 545

Qy 551 ITASQHGMFVSGEAYSASSAMRLKSVYRALLLNSOTLLYDHDIERQEDSPINSVSA 610
Db 536 ITAAQHQDRMPVSGEAYSASSAMRLKSVYRALLLNSOTLLYDHDIERQETSPINSVSA 595

Qy 611 FFINLDIDKIPYKFMNRYNAMMDWDHYKMFWDHHGNSPVAQEAQAEFKKR 670
Db 596 FFINLDIDKIPYKFMNRYNAMMDWDHYKMFWDHHGNSPVAQEAQAEFKKR 655

Qy 671 WTQFVNNTFQESTITRATIAYVFQPYINVSSCRFDISSNPGLOISLNVNTTEHVSVYTD 730
Db 656 WTQFVNNTFHMESTITRATIAYVFQPYINVSSCRFDISSNPGLOISLNVNTTEHVSVYTD 715

Qy 731 YANLKTRENYLGGFFASVADQCOITRQITRGLGTOIAIVKPVRHDRITIPFGFKENIAVGLIL 790
Db 716 YONLKSRFSYLGFFGPAVANQCOITRQITRGLGTOBIVNPVRHDCKVNPFQGFKENIAVGFIL 775

Qy 791 CISLVLVLTQWRWYFLSFRKLMLRMVLILVIALWFLDWSQSOPICAKWTRTEAGSK 850
Db 776 CISLVLVLTQWRWYFLSFRKLMLRMVLILVIALWFLDWSQSOPICAKWTRTEAKANE 835

Qy 851 KSLSEGHMHMDLPDVVITSLPGSGAEILKQLFNPNSDFLFLYRIVTAYIDIPETLEIDSF 910
Db 836 KWNISEGHAVDLPNVITSLPGSGAEILKQLFNPNSDFLFLYRIVTAYIDIPETLEIDSF 895

Qy 911 VDACEWKYSDIRSCHFHLLRGWLSVQDTKLHLQNTHLHEPNRGKLAQYFAMNKDKCRK 970
Db 896 VDACEWKYSDIRSCHFHLLRGWLSVQDTKLHLQNTHLHETRSKLAQYFTMKDKCRK 955

Qy 971 FKRESLPEQRSONKGAFDRDAEVIRALRRHLVYYPSARPVLSSGSVTLKHFPEVYL 1030
Db 956 LRRESLQDQSRSPKGPFDAYIRALRRHLVYYPSARPVLSSGSVTLKHFPEVYL 1015

Qy 1031 GASMRALYTVRDPRAWIYSMLNSKPSLSSLKVNPEHLAKLFKIEGKOKCNCNLSGYAFE 1090
Db 1016 GTSNRALYTVRDPRAWIYSVLSKPSLSSLKVNPEHLAKLFKIEGKOKCNCNLSGYAFE 1075

Qy 1091 YEPURKELSKSKSNASLISHIILANTAAALRINTDLPLPTSQYLVKFEDIVHFPOKTTER 1150
Db 1076 YESULKKELEISQSNASLISHIILANTAAALRINTDLPLPTNHYLVKFEDIVHFPOKTTER 1135

Qy 1151 IFAPLGIPLIPSPASLNQLPATSTNLFLPYEGEISPTNTNWVKNLPROBKILKENICWT 1210
Db 1136 IFAPLGIPLIPSPASLNQMLPATSTNLFLPYEGEISPTNTNWVKNLPROBKILKENICWT 1195

Qy 1211 LMDRGYPKFKMD 1222
Db 1196 LMDHLYPKFKMD 1207

RESULT 6
US-10-697-828-9
; Sequence 9, Application US/10697828
; Publication No. US20040185546A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven
; APPLICANT: Lee, Jin Kyu
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: Novel Glycosyl Sulfotransferases GST-4alpha, GST-4beta, & GST-6
; FILE REFERENCE: UCAL-13801V
; CURRENT APPLICATION NUMBER: US/10/697,828
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: 09-1593,828
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60-144,694
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 596

; TYPE: PRT ; ORGANISM: homo sapiens
 US-10-697-828-9

Query Match 48.2%; Score 3147; DB 16; Length 596;
 Best Local Similarity 100.0%; Pred. No. 5.1e-275;
 Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 627 MNRNGAMMDVDAHYKQFWEDHGNPSMASIOBAAEFKKRWTQFVNNTFQMFSTIT 686
 Db 1 MNRNGAMMDVDAHYKQFWEDHGNPSMASIOBAAEFKKRWTQFVNNTFQMFSTIT 60

Qy 687 RIAVYFCGPINVSSCRIPTDSSNPGLQTSVLNTNTEHVSIITDYNLKTTRPNYLQFGGF 746
 Db 61 RIAVYFCGPINVSSCRIPTDSSNPGLQTSVLNTNTEHVSIITDYNLKTTRPNYLQFGGF 120

Qy 747 ASVADQGQTTRFLGTQAIKVPRVHDRLIFPFGFKENTAVGLICLTSVLTFQMFYFLS 806
 Db 121 ASVADQGQTTRFLGTQAIKVPRVHDRLIFPFGFKENTAVGLICLTSVLTFQMFYFLS 180

Qy 807 FRKLMRWLLVLTALWFELLDWSTCSQPICAKWTRTEAEGSKKSLSEGHMQLDPDV 866
 Db 181 FRKLMRWLLVLTALWFELLDWSTCSQPICAKWTRTEAEGSKKSLSEGHMQLDPDV 240

Qy 867 ITSLPGSCAELKQLFENSSDFLYIRVPTAYDIPETELEIDSFDACEWVSDIRSGHF 926
 Db 241 ITSLPGSCAELKQLFENSSDFLYIRVPTAYDIPETELEIDSFDACEWVSDIRSGHF 300

Qy 927 RLRRGWLSQLVQDPTKLHQINHHIHEPNRGKLAQYFANKDKCRKFKERESPEQRQMKG 986
 Db 301 RLRRGWLSQLVQDPTKLHQINHHIHEPNRGKLAQYFANKDKCRKFKERESPEQRQMKG 360

Qy 987 AFDRDAEYTRALRHLVYVPPSAPVPLSLSGSWTLKLUFFFQEVLGASMRLAYIVDRPAW 1046
 Db 361 AFDRDAEYTRALRHLVYVPPSAPVPLSLSGSWTLKLUFFFQEVLGASMRLAYIVDRPAW 420

Qy 1047 IYSMLYNKSPKPSLKNVPEHLAKLFLEGCKCKCNLNSGAFFEYEPRLKELSKSKSNAY 1106
 Db 421 IYSMLYNKSPKPSLKNVPEHLAKLFLEGCKCKCNLNSGAFFEYEPRLKELSKSKSNAY 480

Qy 1107 SLSLHMLANTAAALRINTDLPLPTSQLVKPFEDIVHPFKQKITERIFAFGLGPLSPASLNQ 1166
 Db 481 SLSLHMLANTAAALRINTDLPLPTSQLVKPFEDIVHPFKQKITERIFAFGLGPLSPASLNQ 540

Qy 1167 ILPATSTNLFLYFPEGEISPTNTNVVKONLPRDEIKLJENICWTLMDRLGPKFMD 1222
 Db 541 ILPATSTNLFLYFPEGEISPTNTNVVKONLPRDEIKLJENICWTLMDRLGPKFMD 596

RESULT 7
 US-09-984-429-71

; Sequence 71, Application US/0984429
 ; Publication No. US20040010132A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 53 Human Secreted Proteins
 ; FILE REFERENCE: P2018P2
 ; CURRENT APPLICATION NUMBER: US/09/984,429
 ; CURRENT FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: 60/061,463
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,591
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/061,143
 ; PRIOR FILING DATE: 1999-04-08
 ; PRIOR APPLICATION NUMBER: PCT/US98/21142
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,463
 ; PRIOR FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: US20030078386A1
 ; GENERAL INFORMATION: Secreted Protein HPEAD48
 ; TITLE OF INVENTION: Secreted Protein HPEAD48
 ; FILE REFERENCE: P2018P1D1
 ; CURRENT APPLICATION NUMBER: US/10/150,111
 ; CURRENT FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,527
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,529
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/071,498
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,532
 ; PRIOR FILING DATE: 1997-10-09
 ; NUMBER OF SEQ ID NOS: 219
 ; SEQ ID NO: 71
 ; LENGTH: 126
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (126)
 ; OTHER INFORMATION: Xaa equals stop translation
 ; US-10-150-11-71

Query Match 8.7%; Score 568; DB 14; Length 126;
 Best Local Similarity 100.0%; Pred. No. 9.1e-13;
 Matches 0; Mismatches 0; Indels 0; Gaps 0;

Qy 811 MRWILLVIALWFELLDWSTCSQPICAKWTRTEAEGSKKSLSEGHMQLDPDVITS 870
 Db 1 MRWILLVIALWFELLDWSTCSQPICAKWTRTEAEGSKKSLSEGHMQLDPDVITS 60

Qy 871 PGSGAELKQLFNFNSSDFLYIRVPTAYDIPETELEIDSFDACEW 917

Db 61 PGSGAEILKQLEFFNSSDFLYIRVPTAYIDIPETELBIDSFVDACEWK 107

; Sequence 97, Application US/10211462

; Publication No. US20040033495A1

RESULT 9

US-10-021-660-128

Sequence 128, Application US/10021660

Publication No. US20030152926A1

GENERAL INFORMATION:

APPLICANT: Murray, Richard

APPLICANT: Glynne, Richard

APPLICANT: Watson, Susan R.

APPLICANT: EOS Biotechnology, Inc.

TITLE OF INVENTION: No. US20030152926A1; Methods of Diagnosis of Angiogenesis, Compositions and Methods of Screening for Angiogenesis

TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis

TITLE OF INVENTION: Modulators

FILE REFERENCE: 018501-0007101S

CURRENT FILING DATE: 2001-10/021, 660

PRIOR APPLICATION NUMBER: US-10-12-06

PRIOR FILING DATE: 2001-02-14

PRIOR APPLICATION NUMBER: US/09/784,356

PRIOR FILING DATE: 2001-02-14

PRIOR APPLICATION NUMBER: US 09/791,390

PRIOR FILING DATE: 2001-02-22

PRIOR APPLICATION NUMBER: US 60/310,025

PRIOR FILING DATE: 2001-08-03

PRIOR APPLICATION NUMBER: US 60/334,244

PRIOR FILING DATE: 2001-11-29

NUMBER OF SEQ ID NOS: 230

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO: 97

LENGTH: 411

TYPE: PRT

ORGANISM: Homo sapiens

US-10-211-462-97

Query Match 2.3%; Score 147; DB 14; Length 411;

Best Local Similarity 18.2%; Pred. No. 0.0008;

Matches 89; Conservative 56; Mismatches 149; Indels 194; Gaps 16;

Qy 813 WILIVLIALMFIEL-----LDWWSCT-----SOPICAKWTRTEAGSKKSLSS 855

Db 5 WKAVLLALASIAIQTAIRTPAKSFHTCPGLAEAGLARLC----BESPTFAYNLSR 59

Qy 856 EGHHMMDLPDVVITSLPGSGAEILKQLEFFNSSDFLYIRVPTAYIDIPETELEIDSFVDACE 915

Db 60 KTH-----ILLATTRGSFSYVQLFVQHLDVFYLFBPL----- 93

Db 60 KTH-----ILLATTRGSFSYVQLFVQHLDVFYLFBPL----- 93

Qy 916 WKVSDIRSGHFLRLRGWLQSLVQDPTKLHLQNTLHLHEPNRQGLAQYFAMNKDKKKFKKRE 975

Db 94 -----YHVNQNTLIPRFQG----- 115

Db 976 SLPPEQRSOMKGAFDRDAEI----- 1013

Qy 1014 LSSGSWTLKHFQEVLGA----- 1042

Db 116 MLGASDLRSLSLYDCDYFLENYIKPPPVNHTTDRIFRGASRVLCSPVCDPFGADLV 175

Qy 1103 SNAVS-----LISHMLANTAAALRINTDLPSTSOLVKFEDVHFPOSTTEIIFAFGJPL 1159

Db 1043 PRAWIYSMLNSKPSLISLNVPEHLAKFIEGGKGKCNLNSGYAFEYPLRKELSKSK 1102

Qy 1160 -----SPASLNQILPATSTNLFLPYEGEISPTNTNWQNLPDEIKLIENT 1207

Db 1103 SNAVS-----LISHMLANTAAALRINTDLPSTSOLVKFEDVHFPOSTTEIIFAFGJPL 1159

Db 280 SNSVSTGLMRPPWLKG-----KMLVRYEDLARNPMKKTTEIYGFGLPILD 325

Qy 1208 CWTLMDRLGX 1217

Db 326 SHVARWIQNTRGPPTLGSHKYGTVRN-----SAATAEKWRFRLSYDIVAFQAQNA 375

Qy 1208 CWTLMDRLGX 1217

Db 376 CQQVLAQLGX 385

RESULT 11

US-10-408-765A-395

; Sequence 395, Application US/10408765A

RESULT 10

US-10-211-462-97

; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fatty, Boin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: TARGETS FOR THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088 465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 395
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-395

Query Match 2.3%; Score 147; DB 16; Length 411;
Best Local Similarity 18.2%; Pred. No. 0.0008; Indels 194; Gaps 16;
Matches 89; Conservative 58; Mismatches 149; Delns 194; Gaps 16;

Query 813 WILIVIAWIEL - LDWSTC - SPICAKWTRTEAGSKSLS 855
Db 60 KTH --- IIIATRSGSSPFVGQFNQHLDVPEFPL - - - - -
Db 94 - - - - - YHVQNTLIPFTQG - - - - - KSPADDRV 115
Db 94 - - - - - YHVQNTLIPFTQG - - - - - KSPADDRV 115

Query 813 WILIVIAWIEL - LDWSTC - SPICAKWTRTEAGSKSLS 855
Db 60 KTH --- IIIATRSGSSPFVGQFNQHLDVPEFPL - - - - -
Db 94 - - - - - YHVQNTLIPFTQG - - - - - KSPADDRV 115

Query 856 EGHHMDLPDVVITSLSGSABEILKQFNSDFLYRVPTAYIDFETEIDSFTVACE 915
Db 60 KTH --- IIIATRSGSSPFVGQFNQHLDVPEFPL - - - - -
Db 94 - - - - - YHVQNTLIPFTQG - - - - - KSPADDRV 115

Query 916 WKVSDIRSGHFRLLRGWLQSLVQDTKLHLONTHHEPNRGKLAQYFAMNKDKRKPKRRE 975
Db 60 KTH --- IIIATRSGSSPFVGQFNQHLDVPEFPL - - - - -
Db 94 - - - - - YHVQNTLIPFTQG - - - - - KSPADDRV 115

Query 976 SLEPQSOMKGAFDRDAEIY - - - - - RALRHLVYVPSARV - - - - - LS 1013
Db 116 MLGASDRLLRSLYDCDLYLENYIKPPVNHHTDRIFRGRASVLCSPVCDPPGPADLV 175
Db 116 MLGASDRLLRSLYDCDLYLENYIKPPVNHHTDRIFRGRASVLCSPVCDPPGPADLV 175

Query 916 WKVSDIRSGHFRLLRGWLQSLVQDTKLHLONTHHEPNRGKLAQYFAMNKDKRKPKRRE 975
Db 60 KTH --- IIIATRSGSSPFVGQFNQHLDVPEFPL - - - - -
Db 94 - - - - - YHVQNTLIPFTQG - - - - - KSPADDRV 115

Query 976 SLEPQSOMKGAFDRDAEIY - - - - - RALRHLVYVPSARV - - - - - LS 1013
Db 116 MLGASDRLLRSLYDCDLYLENYIKPPVNHHTDRIFRGRASVLCSPVCDPPGPADLV 175
Db 116 MLGASDRLLRSLYDCDLYLENYIKPPVNHHTDRIFRGRASVLCSPVCDPPGPADLV 175

Query 1043 PRWYTSMLNSKPSLSSLYKVNPEHLAKLQVFDIVHPEQKTTERFAFLGPL - - - - -
Db 236 PRGLASRSETPRDTY - - - - - RLWRLWIGTGR - - - - - KPNLDTQL-TTCEDF 279
Db 236 PRGLASRSETPRDTY - - - - - RLWRLWIGTGR - - - - - KPNLDTQL-TTCEDF 279

Query 1103 SNAVS - LISHLWILANTAAALRINTLILPSQVLFEDIVHPEQKTTERFAFLGPL - - - - -
Db 280 SNSVSTCLMRPEWLGK - - - - - KMLMRYEDLARNPMKCTEEYGFLGIPLD 325
Db 280 SNSVSTCLMRPEWLGK - - - - - KMLMRYEDLARNPMKCTEEYGFLGIPLD 325

Query 1160 - - - - - SPASINQNLFATSTSNLFLPYEGGISPINTNVNQLPDEIKJENI 1.207
Db 326 SHVARLQNNTRGDPPFLGKHYGTVRN - - - - - SAATAEKWRFLSYDIVAFQNA 375
Db 376 CQQLAQLOGY 385

RESULT 13
US-09-817-514-A-8
; Sequence 8, Application US/09817514A
; Patent No. US20020078478A1
; GENERAL INFORMATION:
; APPLICANT: French-Constant, Richard
; APPLICANT: Rocheleau, Thomas
; APPLICANT: Bowen, David
; APPLICANT: Waterfield, Nicholas
; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: 61645
; CURRENT APPLICATION NUMBER: US/09/817,514
; CURRENT FILING DATE: 2000-03-26
; PRIORITY APPLICATION NUMBER: US 60/191806
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 8
; LENGTH: 411

RESULT 12
US-10-723-866-1544
; Sequence 1544, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Azi, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions & Methods for Screening for Soft Tissue Sarcoma Modulators

; TYPE: PRT
; ORGANISM: Photorhabdus luminescens
; US-09-817-514A-8

Query Match Score 145; DB 9; Length 2504;
Best Local Similarity 18.1%; Pred. No. 0.024;
Matches 231; Conservative 180; Mismatches 443; Indels 422; Gaps 60;

29 FSTFEEVSNSYSEWA-----VFTDIDQPFKTKVQDFRPNQKLLKSMGLPSPLYED 78
Db 1003 PTDWERYNKRKYSTWAGVSELYYPENYVDTQICOTKMD-ALLQSINQLSQLNADTVE 1061
Qy 79 AGEIQAMROKSRA SHLFLPFAIRSAVTY-----MLSNPYIYLDPPKHAD----D 122
Db 1062 AFKTY-LTSPFQVANIKVIAHDNYNVDO8LTYFIGDQAARGTYIYWSVDHRCENSEK 1120
Qy 123 PAA-----KNEIYGNLNLPPALYCLICPEKVKAFEFVLEMDRKYGYKDWLIVENAPGDEV 178
Db 1121 FAANAWGENKIK-----TCAVNPWMKNTIRPVV--YMSRL--YUMLWEQQSKKSD- 1165
Qy 179 PIGHSLTGFAATAFDLFYLNLDNHREOKYLBKIW-----VITEENYEYSVRSMGKQLLH 232
Db 1166 -----DGKTIIYQ--YNLLKLAHIR--YQGSWNTPPTFDTKVNNTSSTDAAESL-- 1212
Qy 233 NHQATNMIAILITGALVGTGVDKGSKANIWKQAVDVMEKTMFLNHHIVDGSSLDESVAYGSY 292
Db 1213 -----GLYCTGYQ-----EDTLHMFYMSOSS-----YSSY 1239
Qy 293 T---AKSVTQVFLAORHFNINNL-----NNWLJRKMHFWFYATTLPSPORTVGIADSNNWF 347
Db 1240 TDNNAPVFTGLYIFADMSSDMNTNAQATNTYNNSY-----PQFDTYMADPSD---- 1286
Qy 348 YGPESQVLVFIDKFILNGNAGNWLAQOIRKRPKGPMVPASTAORMSTLHTEINYDPLQT 407
Db 1287 -----NKVTRVNRYYAEYD-----IPSV-----TSNSNSYMGDHSLT 1323
Qy 408 QPPADYGTAKHTFPNWGVYTYGAGLPN-----TQTNTFVS-FPKSGKLGGRAY 455
Db 1324 -----MLYGGSVPTNITFESAARDLRLSTMNSTIHNQAGTREI 1363
Qy 456 YD1VHFQPSWIDGWRSFNPSPHEHPDQNSTFPAPNGQVFSYALGPKLSHLNNTLVFAP 515
Db 1364 QCNLMMQYASLGDKFLYDSSFD---DANRFLV-----KKGKDENSDDSTCIVNE 1414
Qy 516 SPSSOCNKPWEGSQLGECAQWIKWTGE----EVGDAGEITITASQHGMEMVFSGEAVSAYS 571
Db 1415 NPSS-E DKWYFSSKDNDKNTADYNGTQCAGTSNKDFYTNLQETEVISUTGGYWSYK 1473
Qy 572 SAMRLKSVYALLIANSQTLVVVDHERQDSPINSVSAFPHNLDIEKVIKYKEMNRYN 631
Db 1474 IS-----NPNT-----NTGIDSARV--KVTVKAG 1496
Qy 632 GAMMDVDAHYKMFWDHGNSPMASIQEAQAFKKRNTQFVNNTFQEST--ITRA 689
Db 1497 G-----DQDIFTADNSTVPQ-----QOPASPEEMIYQFNNNTIDCKNLNFIDNQ 1542
Qy 690 YVFYGPYINVSSCRFDSSPGLQISLNVNNTHEVSVIYTVDYHNL-----KTRFNYL 741
Db 1543 HIEIDFTATAQDGFLGAETFLIPVTKVQVLTENVIALYSNGVYQMQGAYRTRLNTL 1602
Qy 742 GFGGFASVADQG-----QITRGEGTOAVKPVYHDRLLIFPGFKENIAGLI 789
Db 1603 FAQOLYVSRAHQDIDAVLSMETONIQBPLQGAGTYVQLVLDYKDESIHGTNSKFAEYVDI 1662
Qy 790 LCISLVLILTFOURFYLUSFRUMRWLILTWIAWFIELL----DWSSTCSOPICAKW--- 841
Db 1663 FKENDSPVYIYQGELSETSQTVVK-----VFLSYFIBATGKQHNL-----VRAVYQKET 1711
Qy 842 -----TRTEA-----GSKKSISSEGHMIDLPDPVITSLPGSCA--- 875
Db 1712 TDKLFLDRTDKEPDGHMFLSDDKTFSGLSSAQALKNDSPEMDF-----SGANAL 1761
Qy 876 -----EILKQLOFFNNSD - FLYIRVPTAYIDIPETELEIDSFVDAECWK 918

Db 1762 YFMBLFYTTPMOMMAHRLLQEONFDAAHWMPRYVWSPSGY-----VDGKTAIYHNNV 1813
Qy 919 SDIERSGHFPLLRQSLQSIYQDFTKLHQNTHLHEPN-----RGKQLOYFA-----MN 964
Db 1814 -----RPEBEDTSWNAQDSDTDEVAQODDPHYKATMFATDLMLAR 1858
Qy 965 KDKKRKFKRKRESLPEQRQSOMKGAFDRDAEYIRALR-----RHUVYPPSARPVLSLSSGS 1018
Db 1859 GDAAYRQLERDTLAEAKW-----YTQALNLLGDEPQWMLSTWANPTLGNAASK 1908
Qy 1019 WTLKLHFFPQEVLGASMALYIVDPRATVSMLYNSKSLYNSKVNPEHLAKLFKIEGK 1078
Db 1909 TTQQVR-----QVQL-TQLRNLNSRVKTP-----LIGTANSLTAL-FLPQENSKL-----K 1951
Qy 1079 G-----KCUNISGAYFEYEPKREIUSKSNSKNAVSLSHLMLANTAAALRINTDLIPT 1130
Db 1952 GYWRFTAQMFMNJRHLNSL-----DGQPLSLPLYAKADPKALLSAAVSASQGSA----- 2001
Qy 1131 SYQLVKFEDIVH-PPQ 1145
Db 2002 --DLPKAPLTIHRPQ 2015

RESULT 14
US-10-262-794A-12 ; Sequence 12, Application US/10262794A
; Publication No. US20030207806A1
; GENERAL INFORMATION:
; APPLICANT: Bnsign, Jerald C
; APPLICANT: Bowen, David J
; APPLICANT: Peter L, James
; APPLICANT: Patig, Raymond
; APPLICANT: Schoonover, Sue
; APPLICANT: ffrench-Constant, Richard
; APPLICANT: Rocheleau, Thomas A.
; APPLICANT: Blackburn, Michael B.
; APPLICANT: Hey, Timothy D.
; APPLICANT: Merlo, Donald J.
; APPLICANT: Orr, Gregory L.
; APPLICANT: Roberts, Jean L.
; APPLICANT: Strickland, James A.
; APPLICANT: Guo, Lining
; APPLICANT: Ciche, Todd A.
; APPLICANT: Sukhapinda, Kititri
; TITLE OF INVENTION: Insecticidal Protein Toxins From Photorhabdus
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow AgroSciences Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/262-794A
; FILING DATE: 02-OCT-2002
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,567
; FILING DATE: 05-MAY-1997
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/063,615
; FILING DATE: 18-MAY-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/395,497
; FILING DATE: 28-FEB-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,255
FILING DATE: 06-NOV-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,423
FILING DATE: 28-FEB-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/705,484
FILING DATE: 28-AUG-1996

ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 960296.93804
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-551-5000
TELEFAX: 608 251-9166

SEQUENCE CHARACTERISTICS:
LENGTH: 2504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-10-262-794A-12

Query Match 2.2%; Score 145; DB 15; Length 2504;
Best Local Similarity 18.1%; Pred. No. 0.024;
Matches 231; Conservative 180; Mismatches 443; Indels 422; Gaps 60;

Qy 29 FSTFEEVSNSYEWA-----VFTDDIDQFKTQRQDFERPNQKLUKKSMLHPSLYF 78
Db 1003 FTDWERYNKYSTNWAGVSELVYYPENYVDPDTQRIGOTKMD-ALLOSIQSQNADIVD 1061
Qy 79 AGELOAMRQKRASHLHLHPLAIRSAVTY----MSNPNTPYVLPPLPQHA----D 122
Db 1062 AFKTY-LTSSQEVALKVTSAYHDNVNTDQGLTYFIGIDQAAIGTYWRSVDISKCENGK 1120
Qy 123 PAA---KNNIEIYGNLNPPLALYCLCPEDKVAFEFVLEYMDRMVGKYDKWLYENAPGDEV 178
Db 1121 FAANAWGENKKI-----TCAVNPMPKNIIRPVV--YMSRL-YUWLWEQSKSD-
Qy 179 PIGHSLTGFTADEFPLYNLIDNHRRQKYLKEIW-----VITEEMYEYSKVRSWKGKOLLH 232
Db 1166 -----DGKTIYQ--YNLKLAIHQ--YDGSNNTPTFEDVTEKVKNYNTSDDAESL-- 1212
Qy 233 NHOATNMIAILITGALVGTGVDGSKANIWKOAVDVMEKTMLFLNHHIVDGSLEDEGVAYSY 292
Db 1213 -----GLYCTGYCQ-----EDTLVNFYMSOSS-----YSSY 1239

Qy 293 T---AKSVTQYVFLAQRHMINNLD--NNWLNKGHWFWYATALLPGQFTVGADSNNW 347
Db 1240 TDNNAPVTGIFTYPAIDMSMTNAQATHYWNNSY-----PQFDYMAFDSD--- 1286
Qy 348 YGPESQLVFFDIFLUNGAGNQIKRKPDKGPVMPSSTAORMSLHETIWYDQLT 407
Db 1287 -----NKKVITRRVNRYAEDYE---IPSSV---TSNSNYSGWDHSLT 1323
Qy 408 ROPPADYGTAKHTPPNGVTVYGAFLN-----TOTNTFYS-FKSGKLGRAY 455
Db 1324 -----MLYCGSVNPNTFESAEDLRLSTNMAISLIHNGYAGTRRI 1363

Qy 456 YDIVHFQPYSWIDGWRSFENGEHPDONSFTFAFANGQVFSEALYGPKLSHANNVLVAP 515
Db 1364 QCNUMKQAYLGDXKEPIYDSSFD--DAHRFLNLP---LF---KFGKDENSISCIYNE 1414
Qy 516 SPSSQCNKPJBQGQJCECAQWLKWTGE---EVGDAAGELTITASQHGMENVFSGEAVSAYS 571
Db 1415 NPSE-DKKHYFESSKDNDKTAIDYNGTQCIAGTSNKDFYNUQEIIVSTVCGWNSSYK 1473
Qy 572 SAMRLKSVVYHALLLNSOTLWVVDHIERQDSPINSVSASFHNLIDPFKXIPKFNMRYN 631
Db 1474 IS-----NPJN1-----NTGIDSAAK--KVTVKAG 1496
Qy 632 GAMMDVWDHYKMFEDHIGNSPMASIQEAEQAAEFKKRWTQFVNVTQMFEST--ITRIA 689

Db 1497 G-----DQGIFTADNSTVVPQ-----OPAPSPEMIXQFNNTIDNQ 1542
Qy 690 YFYGPYINVSSCRFIDSNPGQLSUNVNTTEHVSYTVDYHNL-----KTRFNYL 741
Db 1543 HIEIDFTATAQDGREFLGAETPFLIVPKVKGTVNTIALYSENNGYQMQGAYRTRLNL 1602
Qy 742 GFGGFASVADQG-----QTRFGLGTOAIVKPVRHDLIIPFGPKENIAVGLI 789
Db 1603 FAQQLVSRANGIDAVISMETONIQEPQLGAGTYVQLVLDKYDESIGHTNKSFALEYVDI 1662
Qy 790 LCISLVLTFOWRFYLSFRKLMRLLVIALWFIELL----DVRSTSQSPICAKW---- 841
Db 1663 FKENDSSVIYQELSETSQPTVVK----VFLSYFTEATGNKNHLW----VRAKYQKET 1711
Qy 842 -----TRTEAE-----CSKKSLSSSEGHMFLDVVITSLPGSGA--- 875
Db 1712 TDKILFORTDEKDPHGWLFSDDDHKTFSGLSSAQALKNDSEBMDF----SGANAL 1761
Qy 876 -----BILQOLFENSSD--FLYIRVBTAYDIPETELETIDSFDACEWKV 918
Db 1762 YFWELFLYYPMMAMHRLQEQNFDAANHWRFYWSPSGYI----VDGKIAIYHNV 1813
Qy 919 SDIRSGHFRLLRGWLSQYDQTKLHQINTLHEPN-----RGKLAQYFA----MN 964
Db 1814 -----RPLEEDTSWAAQQLDSTDPAVAQDDPMHYKVAWMATLDILMAR 1858
Qy 965 XDKKRKFRKRSILPEQRSQMGAFDRAEYTRAIL----RHLYYPPSARPVLSSGS 1018
Db 1859 GDAAYRQLERDTLAEARWW-----YTOQALNLGDEQVMILSTWANTPLGNAASK 1908
Qy 1019 WTLKLHAFQEYVGAHSHRALYTVRDPAWYIYMSLYNKSPLSYLSKRNVPHELAKLPKTIEGGK 1078
Db 1909 TTQQR--QQLV-TQLRNLNSRKTP----LLGTANSLTAL-FLPQENSKL----K 1951
Qy 1079 G-----KCNLNNSCYA邢EYBPLRKLSELSKSNSAVSLLSHMLANTAAALRINTDLPLT 1130
Db 1952 GYWRTLAQMFMNLRNLNSIDQPLSLPYAKPADPKLSSAAVSAQSQGA----- 2001

RESULT 15
US-10-754-115-59
Sequence 59, Application US/10754115
Publication No. US20040208907A1
GENERAL INFORMATION:
APPLICANT: Schleifer, Amanda
APPLICANT: Bevan, Scott
APPLICANT: Bintrim, Jon
APPLICANT: Mitchell, Jon
APPLICANT: Li, Ze Sheng
APPLICANT: Ni, Weiting
APPLICANT: Zhu, Baolong
APPLICANT: Merlo, Don
APPLICANT: Apel-Hirkhold, Patricia
APPLICANT: Meade, Thomas
TITLE OF INVENTION: Mixing and Matching TC Proteins for Pest Control
FILE REFERENCE: DRS-104XCI
CURRENT APPLICATION NUMBER: US/10/754,115
CURRENT FILING DATE: 2004-01-07
PRIOR APPLICATION NUMBER: US 60/441,723
PRIOR FILING DATE: 2003-01-21
NUMBER OF SEQ ID NOS: 64
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 59
LENGTH: 2504
TYPE: PRT
ORGANISM: Photorthabdus luminescens
US-10-754-115-59

| | | |
|--|--|--|
| Query Match Matches 231; Conservative 180; Mismatches 443; Indels 422; Gaps 60; | Score 145; DB 16; Length 2504; | Qy 919 SDIRSGHFRLRGWLOSLVQDTKLHLNTHHEPN-----RGKLAQYFA-----MN 964 |
| Best Local Similarity 18.1%; Pred. No. 0.024; | Db 1814 -----RPEBEDSWNAQDSTDPAVAQDDPMHYKATFMTDLIMAR 1858 | Db 1814 -----RPEBEDSWNAQDSTDPAVAQDDPMHYKATFMTDLIMAR 1858 |
| Qy 29 FSTFFEBVSNTSEWA-----VFTDDIDQFXTQVQDFRNQKLUKKSMLHPSLYFD 78 | Qy 965 KDKKRKFKRESLPEQRSQMKGAFDRDAEYIRALR-----RHLVYPPSARPVLSSGS 1018 | Qy 965 KDKKRKFKRESLPEQRSQMKGAFDRDAEYIRALR-----RHLVYPPSARPVLSSGS 1018 |
| Db 1003 FTDWERYNKYSTWAVSELVYPPENYVDETOIGQTMRMD-ALLQSINQSQLNADTYED 1061 | Db 1859 GDAYAYROLERDTLAEAKW-----YTOQNLNLGDEPOVMSLSTWANPTLGNAASK 1908 | Db 1859 GDAYAYROLERDTLAEAKW-----YTOQNLNLGDEPOVMSLSTWANPTLGNAASK 1908 |
| Qy 79 AGEIQAMRQSRASHLHPLFAIRSAYTV-----MLSNTYLLPPPKHA-----D 122 | Qy 1019 WTURKLHFFQEVLGASMRALYIVRDPRAWIYSMAYNSKESPLSYLSKVNPHLAKLFKIEGK 1078 | Qy 1019 WTURKLHFFQEVLGASMRALYIVRDPRAWIYSMAYNSKESPLSYLSKVNPHLAKLFKIEGK 1078 |
| Db 1062 AFKTY-LTSESEQVANLKVTSAYHDNVNVDQGGLTYFIGIDOAQPCYYWSVDSIKCENGK 1120 | Db 1909 TTQQVR-QQVL-TQLRULNSRVKTP-----LLGTANSLTL-ELFOENSKL-----K 1951 | Db 1909 TTQQVR-QQVL-TQLRULNSRVKTP-----LLGTANSLTL-ELFOENSKL-----K 1951 |
| Qy 123 FA-----KNEIYGNLNPPLAIALYCLICLCPETKVAFFEVLEMDRMVGYKDWLIVENAPGDEV 178 | Qy 1079 G-----KCMLNSGKAFYEYPLRKELSKSKSNASVSLISHLWANTAAIRINTDILJPT 1130 | Qy 1079 G-----KCMLNSGKAFYEYPLRKELSKSKSNASVSLISHLWANTAAIRINTDILJPT 1130 |
| Db 1121 PAANAWGENKIK-----TCAVNPNKQNIIRPVV--YMSRL--YLWLEQOSKKSD- 1165 | Db 1952 GYVRTLAORMENLRHNLSDGQLSSLPIYAKPADPKAIIUSAWSAQSGA-----2001 | Db 1952 GYVRTLAORMENLRHNLSDGQLSSLPIYAKPADPKAIIUSAWSAQSGA-----2001 |
| Qy 179 PIGHSL-TGFATAFDFLYNLLDNHRQKYLEKIW-----VITEEMEYYSKVRWSKGKOLH 232 | Qy 1131 SYQLVKFEDIVH-FPQ 1145 | Qy 1131 SYQLVKFEDIVH-FPQ 1145 |
| Db 1166 -----DGKTIYQ--YNLKLAHIR--YGSWNTTPFTDVTKEVNTISSTDAAESL-- 1212 | Db 2002 --DLPKAPLTIRFPQ 2015 | Db 2002 --DLPKAPLTIRFPQ 2015 |
| Qy 233 NHQATNMIAILTGALVTGWDKGSKXANTIKWGAVIDVMEKTMFLLAHIVDGSLSDEGVAYSSY 292 | Search completed: June 23, 2005, 09:28:02 Job time : 159.482 secs | |
| Db 1213 -----GLYCTGYQG-----EDTLYGTYSMQSS-----YESY 1239 | | |
| Qy 293 T-----AKSVTOVFLAQRFNINLD--ANWLMKHPWYVATLTFQGQFQRTVGIADSNTYWF 347 | | |
| Db 1240 TDNNAPVTGUYFADMSSDMDTNAQATNTYNNSY-----PQFDTWMADPSSD--- 1286 | | |
| Qy 348 YGPESQLVFLDKFLKRNAGCNWLAQOQIRKRPKGPMVPSTAQRWSTLHTEYINYDPOLT 407 | | |
| Db 1287 -----NKKVITRRVNNRYAEDY-----IPSSV---TSNSNYSMGDHSLT 1323 | | |
| Qy 408 POPPACYGTAKIHTFPNGWGYTYGAGLPN-----TOTNTFVUS-FKSGKLGGRAV 455 | | |
| Db 1324 -----MLYGSVPTNITFESEAIDLRLSTMNALSTNMLAISTHNGYAGTRRI 1363 | | |
| Qy 456 YDIVHFPQPSWIDGWRSFNGPNSGHERPDQNSTPAPNGQVFTSEALYGPKUSHLNNVLFA 515 | | |
| Db 1364 QCNLMKQYASLGDKFTIYDSSFD--DANRPNLVP---LF---RFGKDENSDDJCICNB 1414 | | |
| Qy 516 SPSSOCNKPWEGLGECAQWLKWTEB-----EVGDAAGEITITASOHGEMYVFSGBAVSAYS 571 | | |
| Db 1415 NSSE-DKKNYFSSKDNNKTAIDYNGTQC-DAGTSNKNDFYNLQIEVISTGGWSSYK 1473 | | |
| Qy 572 SAMRLKSVYALLLNQSTLUVVDIERQBDSPINSVSAFHNLIDFCKTIPYKEMNRYN 631 | | |
| Db 1474 IS-----NPINI-----NTCIDSARK--KVTKVAG 1496 | | |
| Qy 632 GAMMDVWDAKYMFWDHGNSPMASIQUEAAEPKRMWTOFVNVTQBNB-----KTRFENL 689 | | |
| Db 1497 G-----DQDQIFTADNSSTVVPQ-----QAPSPSEEMYQFNNTLIDCKNLFIDNOA 1542 | | |
| Qy 690 YVFYGPYINVSSCRFDSSNPGLOISLNNTNTEHYVSIYTDYHNL-----KTRFENL 741 | | |
| Db 1543 HIEIDFTATAODGRFLGAEFTIPIVTKVQUGTENTIALWMPFELL----DWSTCSOPICA 1602 | | |
| Qy 742 GFGGFASVADOG-----QITRFLGFTQAIKPYRHDRIFFPFKFKENIAGLI 789 | | |
| Db 1603 FAQQLVSRANGRIDAVLSMTEQNIQEPOLGAGTYQVLVLUKYDSTHGNTKSFALEYVDI 16662 | | |
| Qy 790 LCISLVLTFOWRFYLSFRKLMRWWLILVIALWMPFELL----DWSTCSOPICA 841 | | |
| Db 1663 FKENDSFVYQGELSETSQTVVK-----VFTSYFILEATGKONHLM-----VRAYQKET 1711 | | |
| Qy 842 -----TRTEAE-----GSKKSLSSEGHHMDLDPDVVITSLPGSGA 875 | | |
| Db 1712 TDKILFDRTEKDPHGMFLSDDHKTFSGLSSQAQKNDSEPMDF-----SGANAL 1761 | | |
| Qy 876 -----EILKQFLPNSSD-----FLYIYRPTAYDIPEPELETISFVDACEWKY 918 | | |
| Db 1762 YFWELFYTTPMMAHRLQEQNFDAANHWPYVWSPSGYI-----VNGKIAIYHNNV 1813 | | |

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model
Run on: June 23, 2005, 08:33:38 ; Search time 27.7171 Seconds.
(without alignments)
4242.027 Million cell updates/sec

Title: US-10-697-828-15
Perfect score: 6332
Sequence: 1 MPKGGAPPWIMALMFTGHLL.....LIENICWTLMDRGYPKFMD 1222

Scoring table: BLOSUM22
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*
Maximum Match 100*
Listing first 45 summaries

Database : PIR_79;*
1: pir1;*
2: pir2;*
3: pir3;*
4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|----------------------|
| 1 | 251.5 | 3.9 | 672 | 2 T36083 | hypothetical proteo |
| 2 | 192 | 2.9 | 776 | 2 D98354 | oligo alginate lyase |
| 3 | 192 | 2.9 | 776 | 2 AC2928 | oligo alginate lyase |
| 4 | 133 | 2.0 | 484 | 2 JB0261 | N-acetyl glucosamin |
| 5 | 127 | 1.9 | 1456 | 1 A36563 | mannose receptor P |
| 6 | 125 | 1.9 | 4344 | 1 A53489 | dynein heavy chain |
| 7 | 123 | 1.9 | 1044 | 2 B72338 | galactosyltransfer |
| 8 | 122 | 1.9 | 1747 | 2 T43162 | vitellogenin - GYD |
| 9 | 120.5 | 1.8 | 944 | 2 AE2926 | hypothetical proteo |
| 10 | 113.5 | 1.7 | 484 | 2 JC7350 | N-acetyl glucosamin |
| 11 | 113.5 | 1.7 | 3744 | 2 S46715 | hypothetical proteo |
| 12 | 112.5 | 1.7 | 2216 | 2 S78398 | hypothetical proteo |
| 13 | 112.5 | 1.7 | 4199 | 2 S76412 | isoleucyl-tRNA syn |
| 14 | 111 | 1.7 | 921 | 2 AE1698 | mannose receptor P |
| 15 | 110.5 | 1.7 | 1248 | 2 AE9874 | bromodomain protei |
| 16 | 110.5 | 1.7 | 1455 | 1 A48925 | genome polyprotein |
| 17 | 110 | 1.7 | 542 | 2 S54260 | TRAP-like protein |
| 18 | 110 | 1.7 | 2095 | 2 S29529 | hypothetical proteo |
| 19 | 110 | 1.7 | 3695 | 2 T38084 | rhoptry protein - |
| 20 | 109.5 | 1.7 | 1441 | 2 T00335 | probable lipoxygen |
| 21 | 109 | 1.7 | 2401 | 2 T28676 | probable cytochrom |
| 22 | 108.5 | 1.7 | 866 | 2 T06454 | chitinase A / impor |
| 23 | 108.5 | 1.7 | 936 | 2 H71862 | C29E14 protein - |
| 24 | 108 | 1.7 | 1215 | 2 T247916 | hypothetical proteo |
| 25 | 108 | 1.7 | 1102 | 2 S44772 | fatty-acyl-CoA syn |
| 26 | 108 | 1.7 | 1589 | 2 T22668 | hypothetical proteo |
| 27 | 107 | 1.7 | 2076 | 2 S15999 | viral protein - Dic |
| 28 | 107 | 1.6 | 771 | 2 H84845 | amylase A (EC 3.2. |
| 29 | 107 | 1.6 | 1131 | 2 F82875 | .viral protein - Dic |

ALIGNMENTS

| | | |
|---|---|--|
| RESULT 1 | T36083 | hypothetical protein SCB134.01c - Streptomyces coelicolor (fragment) |
| C;Species: | Streptomyces coelicolor | |
| C;Date: | 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-Jul-2004 | |
| C;Accession: | T36083 | |
| R;Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. | | |
| R;Submitted to the EMBL Data Library, April 1999 | | |
| A;Accession: | T36083 | |
| A;Status: preliminary; translated from GB/EMBL/DDBJ | | |
| A;Residues: 1-672 <SAU> | | |
| A;Cross-references: UNIPROT:Q8CTW4; EMBL:AL049661; PIDN:CRB41199.1; GSPDB:GN00070; SCOEDB | | |
| A;Experimental source: strain A3 (2) | | |
| C;Genetics: | | |
| A;Gene: SCB134.01C | | |
| C;Superfamily: Oligosaccharide lyase | | |
| Query Match Score 3.9*; Score 251.5; DB 2; Length 672; | | |
| Best Local Similarity 22.6*; Pred. No. 9.5e-10; | | |
| Matches 125; Conservative 86; Mismatches 249; Indels 93; Gaps 27; | | |
| Qy 80 GETQAMRQKSRAHLHPLAIRSAVTWMSLNPTVYL -- PPPKGADPAKKNQNEYGNNLPI 136 | | |
| Db 6 GRDLSDRLGDLNGTHAAQRRLHEQC----DWYRQNPTPEHPEA 55 | | |
| Qy 137 PLAYLCYKPEDKVAFEEFYEMDRMVGYKDMLVNEAPDEVPIGHSLTGFEATAFDFLYN 196 | | |
| Db 56 NLALAYRUTGQ-RGYLEEARWRWISTCVAYPHMGRAMPDHLDAGWLHLGLSLLAYSWLG 114 | | |
| Qy 197 LLNDHRQKYLEKIWVTEEMEYSKV -- RSNQKQLLHNHOATNMIALTGAIVTGVDK 253 | | |
| Db 115 DLPERPRTILRAXLEQGERLHSFAETTGRWNSAYQNH--NWIC-WTGATAGYAL 170 | | |
| Qy 254 GSXANIKWQAVVDMEKTMFLLAHIVDGSLLDEGVAYGSYTAKSVTQYVFLAQRH----- 307 | | |
| Db 171 GRSE--WTKAARANLETILMLP--DSDSEGUVWRYGVPNLAIHTDLVQREQADLW 226 | | |
| Qy 308 ---FNINNLDNNWKLMMHFWWYATLPGFQRTVGLADSNSNNWFYGPESQLVFLDKFILK 363 | | |
| Db 227 STGCFRLRT--TRW-RLH-----QSAPEFEENIDHGDC-HDRRSGHSVALLYKLASAYQ 276 | | |
| Qy 364 NGGNWLAQOI-RKH-----RPGDPMPVSTAQBWSTLHTEIYWDQPLTP-QPP 411 | | |
| Db 277 DGTQAWLGHNUVAERHFWRAYESGVPGMP-----EAFLLEWYDPRVTPAAPD 326 | | |
| Qy 412 ADYGTAKITHFPNGVWYAGLPNTQNTFVSFKSGKUGLGRAVYDIVFQPWSWIDGPR 471 | | |
| Db 327 REPFTA--YFPDILQQTARTGHDSSA-TCVSKFAAPGGHRAWDEGH-RLKAAGGND 379 | | |
| Qy 472 SFNPGEHEPDQNSFTFPNGQVFSEALYG--PKLSHLNINVLYFAPSPSSQCNKPWEGQL 529 | | |

| | | | | | | | |
|---|--|---|-----|----|-----|--|-----|
| Db | 380 | AMSAGHHHPAGAPVYLHSGAFLAVDEGSNTHKRAAHNLVY-----DGRGWA | 428 | Qy | 502 | PKLSHLNINVLFAPSPSSQCNKPWEGQ-----LGECAOQLKRTGEEVGDAGEBITA | 553 |
| Db | 555 | -----YVAF-----NSQMLNWRQRQTSGNAVLGKQYAEKDALKALARABRIVSV | 602 | Db | 555 | -----YVAF-----NSQMLNWRQRQTSGNAVLGKQYAEKDALKALARABRIVSV | 602 |
| Qy | 530 | GECAGLWLRKG--BEGVDAEGEITIASQHNGEMVFSGEAVASAYSSAMRLKSYTRALLJLN | 587 | Qy | 554 | SOHGENMVFSGEAVASAYSSAMRL-KSVMYRALLNSQTLVVDHTERQE-----601 | |
| Db | 429 | DE-GRYHVYGEPEERRARVRDVL--AQDG-FAHATAESAMFESERLGVQRVDRTLVVTTP | 484 | Db | 603 | BEQPGRHRIVEDATAAXQVANPLVKVLRETHFVNDSYFVIVDEVECSPEOELWLCHLT | 662 |
| Qy | 588 | SQTLVVDHTERQ | 600 | Qy | 602 | DSPINSUASSPHNLIDFKVLPYKEMPRYNGAMMDWDAHYKMFEDHHNSPMSAQI | 659 |
| Db | 485 | LGRVVLDELEAE | 497 | Db | 663 | GAPOTCRSSF-----RYNRKA---GFPQFVYSSGETPQISAVE | 699 |
| RESULT 2 | | | | | | | |
| D98354 | D98354 alginate lyase (AB011415) [imported] - Agrobacterium tumefaciens (strain C58, Cere C;Species: Agrobacterium tumefaciens | | | | | | |
| C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 12-Jul-2004 | C;Accession: AC2928 C;Sequence_revision 11-Jan-2002 #text_change 12-Jul-2004 | | | | | | |
| R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 3232-3238, 2001 | C;Accession: AC2928 C;Title: Genomic Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum | | | | | | |
| A;Reference number: A97359; PMID:21608551; PMID:11743194 | A;Authors: Yoo, H.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, Ester, B.W. | | | | | | |
| A;Accession: D98354 | A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58. | | | | | | |
| A;Status: preliminary | A;Reference number: AB2577; MUID:21608550; PMID:11743193 | | | | | | |
| A;Molecule type: DNA | A;Accession: AC2928 | | | | | | |
| A;Cross-references: UNIPROT:Q8UBJ1; GB:AE007870; PIDN:AAL43841.1; PID:917741383; GSPPDB:G | A;Residues: 1-776 <KTR> | | | | | | |
| C;Genetics: | A;Cross-references: UNIPROT:Q8UBJ1; GB:AE007870; PIDN:AAL43841.1; PID:917741383; GSPPDB:G | | | | | | |
| A;Gene: AGR_L_3558 | A;Experimental source: strain C58 (Dupont) | | | | | | |
| A;MspI position: linear chromosome | C;Genetics: Atu3025 | | | | | | |
| C;Superfamily: oligosaccharide lyase | A;Map position: linear chromosome | | | | | | |
| Query Match 2.9%; Score 192; DB 2; Length 776; | C;Superfamily: oligosaccharide lyase | | | | | | |
| Best Local Similarity 19.1%; Pred. No. 2.e-05; | Query Match 2.9%; Score 192; DB 2; Length 776; | | | | | | |
| Matches 137; Conservative 94; Mismatches 279; Indels 208; Gaps 31; | Best Local Similarity 19.1%; Pred. No. 2.e-05; | | | | | | |
| Qy 33 EESVNSYSEWAVFTDDIDQFKTOKVQDFRPNQKLKSM-----HPSLYFDA 79 | Matches 137; Conservative 94; Mismatches 279; Indels 208; Gaps 31; | | | | | | |
| Db 99 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146 | Qy 33 EESVNSYSEWAVFTDDIDQFKTOKVQDFRPNQKLKSM-----HPSLYFDA 79 | | | | | | |
| Qy 80 GEIQAMRQ--KSRASHLHLFRAIRSAVTWMLSNPNTYLLPP--PKHADFAAKWNEIYGNNL 135 | Db 99 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146 | | | | | | |
| Db 147 EQLSAFADAVAKDPNHCGMAEYKSVFPEWLERPVMPBPQPYENNTNTRVATLWRCMY----202 | Qy 80 GEIQAMRQ--KSRASHLHLFRAIRSAVTWMLSNPNTYLLPP--PKHADFAAKWNEIYGNNL 135 | | | | | | |
| Qy 136 PPLAUYCLQPEDKVAFFKLEDMRMYKD----WLV-----ENAGDE 177 | Db 147 EQLSAFADAVAKDPNHCGMAEYKSVFPEWLERPVMPBPQPYENNTNTRVATLWRCMY----202 | | | | | | |
| Db 203 -----TQQEYVIAIRH-LAAGRVLGRDLDLDAKSRLWAVAADTKGATSRAYNE 254 | Qy 136 PPLAUYCLQPEDKVAFFKLEDMRMYKD----WLV-----ENAGDE 177 | | | | | | |
| Qy 178 VPIGHSLTGTATAFDLYNL--DNHR--RQKYLEKIWITTEEMYEYSKVRSGWKQLLHN 233 | Db 203 -----TQQEYVIAIRH-LAAGRVLGRDLDLDAKSRLWAVAADTKGATSRAYNE 254 | | | | | | |
| Db 255 AGF-RVYVLAQWGYDWLHDLSDEERRTVSLLERPREADVIARIHYFP---YDS 310 | Qy 178 VPIGHSLTGTATAFDLYNL--DNHR--RQKYLEKIWITTEEMYEYSKVRSGWKQLLHN 233 | | | | | | |
| Qy 234 HOATNMIAUUTGALVTDKGSKANIKWQAVDVMEMKTMELINHIVDGSLDE----G 286 | Db 255 AGF-RVYVLAQWGYDWLHDLSDEERRTVSLLERPREADVIARIHYFP---YDS 310 | | | | | | |
| Db 311 HAVRSLSAVLTPTCAIALQESDEAGEWFLYDVTBFL-ATLYSPWAGTDPGMVPSAORWS 393 | Qy 234 HOATNMIAUUTGALVTDKGSKANIKWQAVDVMEMKTMELINHIVDGSLDE----G 286 | | | | | | |
| Qy 287 VAYSYTAKSVTQYV--FLAQRFENINNLDNNVLMKHFVWYATLFCFORTVGIADSNY 344 | Db 311 HAVRSLSAVLTPTCAIALQESDEAGEWFLYDVTBFL-ATLYSPWAGTDPGMVPSAORWS 393 | | | | | | |
| Db 370 MAYLJEANLIRSTIGDLYQRPE-----FQNTGREFLTYKAPGTRRN- 413 | Qy 287 VAYGSYTAKSVTQYV--FLAQRFENINNLDNNVLMKHFVWYATLFCFORTVGIADSNY 344 | | | | | | |
| Qy 345 NWFYGPESOFLDKF-----ILKNGAGNWLAAQIRKXHPKGPMVPSAORWS 393 | Db 370 MAYLJEANLIRSTIGDLYQRPE-----FQNTGREFLTYKAPGTRRN- 413 | | | | | | |
| Db 414 ---FGDDSTLGLDPLGLKLGYNVROFAGTGNGHYQWYDHK-----ADATGTE 459 | Qy 345 NWFYGPESOFLDKF-----ILKNGAGNWLAAQIRKXHPKGPMVPSAORWS 393 | | | | | | |
| Qy 394 TLHTEVYIWYD-----POLTPQPPADYGTAKHTFPNWGVYTFAGLPLNTQTNT 441 | Db 414 ---FGDDSTLGLDPLGLKLGYNVROFAGTGNGHYQWYDHK-----ADATGTE 459 | | | | | | |
| Db 460 MAFTNYGWDLNFDDLTVRHDYPCEAVASPADLPALAVERDGGATIQKDMEDPDRHLQ- | Qy 394 TLHTEVYIWYD-----POLTPQPPADYGTAKHTFPNWGVYTFAGLPLNTQTNT 441 | | | | | | |
| Qy 442 FVSKSGKUGLGRAYDIVFQPSWIDGMRSFNPNGHEHPDONSFTAPNGQVFSEALEYG 501 | Db 460 MAFTNYGWDLNFDDLTVRHDYPCEAVASPADLPALAVERDGGATIQKDMEDPDRHLQ- | | | | | | |
| Db 519 FV-FKS-----SPYGS-----SHSGDQNAFVLYAHGEDLAIQSGY- 554 | Qy 460 MAFTNYGWDLNFDDLTVRHDYPCEAVASPADLPALAVERDGGATIQKDMEDPDRHLQ- | | | | | | |

RESULT 3
AC2928
oligo alginate lyase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Accession: AC2928
C;Sequence: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 12-Jul-2004
C;Accession: AC2928
C;Title: Genomic Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; PMID:21608551; PMID:11743194
A;Accession: D98354
A;Status: preliminary
A;Molecule type: DNA
A;Cross-references: UNIPROT:Q8UBJ1; GB:AE007870; PIDN:AAL43841.1; PID:917741383; GSPPDB:G
A;Gene: Atu3025
A;Map position: linear chromosome
C;Superfamily: oligosaccharide lyase

Query Match 2.9%; Score 192; DB 2; Length 776;
Best Local Similarity 19.1%; Pred. No. 2.e-05;
Matches 137; Conservative 94; Mismatches 279; Indels 208; Gaps 31;
Qy 99 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 99 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 99 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 99 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 100 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 100 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 101 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 101 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 102 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 102 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 103 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 103 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 104 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 104 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 105 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 105 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 106 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 106 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 107 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 107 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 108 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 108 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 109 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 109 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 110 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 110 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 111 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 111 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 112 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 112 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 113 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 113 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 114 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 114 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 115 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 115 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 116 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 116 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 117 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 117 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 118 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 118 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 119 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 119 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 120 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 120 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 121 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 121 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 122 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 122 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 123 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 123 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 124 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 124 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 125 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 125 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 126 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 126 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 127 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 127 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 128 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 128 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 129 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 129 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 130 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 130 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 131 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 131 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 132 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 132 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 133 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 133 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 134 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 134 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 135 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 135 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 136 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 136 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 137 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 137 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 138 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 138 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 139 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 139 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 140 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 140 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 141 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 141 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 142 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 142 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 143 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 143 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 144 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 144 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 145 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 145 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 146 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 146 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 147 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 147 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 148 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 148 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 149 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 149 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 150 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 150 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 151 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 151 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 152 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 152 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 153 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 153 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 154 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 154 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 155 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 155 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 156 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 156 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 157 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 157 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 158 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 158 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 159 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 159 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Qy 160 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146
Db 160 QKSATAHNSWT-----VRSFEISSEALPKTPLPGRSARHAAQTHPRLWLN 146

Qy 442 FVSPFKSGKLUKGRAVYDIVHQPWYDGRVRSFNPQHGHEHPDONSPTFAPNGQVFVSALYG 501
 Db 519 FV_FKS-----SPIGSL-----SHSHGDNQAFVLYAHGDELAQSY- 554
 Qy 502 PKLHLNMYLVFAPSPSSQCNKPWEGO-----LGECAQWLKWTGEEVGDAAGEITTA 553
 Db 555 -----YVAF----NSMQLNVRQRTRSRSSNAVLGGKQYAEKDKAALARAGRIVSV 602
 Qy 554 SGHECMVYFSGEAVSAYSSAAMR1-KS1YVRAILLIINSQTLVYDIERQE- 601
 Db 603 EEPQGHVRIWTFDAPTAQVANPLYQKVIRETHPWNDSFIVUDEVCSBPOELQWLCHTL 662
 Qy 602 DSPINNSVSFAFNLIDDFK1IPYKFMRNRYNGAMDVWDAYKHFWDHGNSPMASIQ 659
 Db 663 GAPQTGRSSF-----RYNGRKA---GFYFGQFVYSSCGTPQISAVE 699

RESULT 4
 JBB261 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
 C;Species: Homo sapiens (man)
 C;Accession: JBB261
 R;Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, Q.W.; Mitsuoka, T.; Biochem. 124, 670-678, 1998
 A;Title: Human N-Acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of
 A;Reference number: JE0261; MUID:98391845; PMID:9722682
 A;Accession: JE0261
 A;Molecule type: mRNA
 A;Residues: 1-484 <UCH>
 A;Cross references: DDBJ:AB014679
 C;Comment: This protein catalyzes the transfer of sulfate from 3'-phosphoadenosine 5'-triphosphate to the 6-OH group of GlcNAc residues.
 C;Superfamily: Chondroitin 6-sulfotransferase
 C;Keywords: sulfotransferase

Query Match 2.0% ; Score 133; DB 2; Length 484;
 Best Local Similarity 26.2%; Pred. No. 0.18;
 Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Db 338 GGPADYHAL-----GMEVICNSMAKTLQALQ-PFDWLQGHLYVRYEDLVGEPV 387
 Qy 1034 MRALYIVDRPRAWIYSMLYNPSKPSLY--SLKVNPE-----HLALKFIEGGK-GKCNLNS 1085
 Db 279 LKVTHLVRDARA-VASSRIRSHGLIRESQVRSRDPRAHRMPFLEAGHKLGAKEGV 337
 Qy 1086 GYAFYEPEPKELSKSKSNVSLSHLMLANTAAALRINTDLPLTSYOLKFEDIYHPQ 1145
 Db 338 KTERIFAFGIPISPA---SLSNOILFATSTNLFLYLYGEIISPTN---TNWQNQNL 1197
 Qy 1146 KTLRRVYDFGLLSPMEMBPAFLNTSGSCSSSKRFV---VSARNATOANAWRTALT 442
 Db 443 FQKQVEEBCYOPHAVLGYER 464

Query Match 2.0% ; Score 133; DB 2; Length 484;
 Best Local Similarity 26.2%; Pred. No. 0.18;
 Matches 53; Conservative 36; Mismatches 81; Indels 32; Gaps 9;

Db 338 GGPADYHAL-----GMEVICNSMAKTLQALQ-PFDWLQGHLYVRYEDLVGEPV 387
 Qy 1034 MRALYIVDRPRAWIYSMLYNPSKPSLY--SLKVNPE-----HLALKFIEGGK-GKCNLNS 1085
 Db 279 LKVTHLVRDARA-VASSRIRSHGLIRESQVRSRDPRAHRMPFLEAGHKLGAKEGV 337
 Qy 1086 GYAFYEPEPKELSKSKSNVSLSHLMLANTAAALRINTDLPLTSYOLKFEDIYHPQ 1145
 Db 338 KTERIFAFGIPISPA---SLSNOILFATSTNLFLYLYGEIISPTN---TNWQNQNL 1197
 Qy 1146 KTLRRVYDFGLLSPMEMBPAFLNTSGSCSSSKRFV---VSARNATOANAWRTALT 442
 Db 443 FQKQVEEBCYOPHAVLGYER 464

RESULT 5
 A36563 mannose receptor precursor - human
 C;Species: Homo sapiens (man)
 C;Date: 10-Sep-1999 #sequence 10-Sep-1999 #ext_change 09-Jul-2004
 C;Accession: A36563; A66926; A44255; B44255; C44255; D44255; F44255; H44
 R;Taylor, M.E.; Conary, J.T.; Lenhardt, M.R.; Stahl, P.D.; Drickamer, K.
 J. Biol. Chem. 265, 12156-12162, 1990
 A;Title: Primary structure of the mannose receptor contains multiple motifs resembling c
 A;Reference number: A36563; MUID:9024192; PMID:237365
 A;Accession: A36563
 A;Molecule type: mRNA
 A;Residues: 1-1456 <TA>
 A;Cross-references: UNIPROT: P228997; GB:J05550; NID:g188675; PID:9188676

Query Match 1.9% ; Score 127; DB 1; Length 1456;
 Best Local Similarity 18.7%; Pred. No. 2.7;
 Matches 138; Conservative 97; Mismatches 284; Indels 218; Gaps 36;

Db 518 IGHSLTGATATAFELYN-----LLDNRARQKYLEKIKIVITEEMEY---SKYRSWG--- 227
 Qy 180 IGHSLTGATATAFELYN-----LLDNRARQKYLEKIKIVITEEMEY---SKYRSWG--- 227
 Db 518 IGHSLTGATATAFELYN-----LLDNRARQKYLEKIKIVITEEMEY---SKYRSWG--- 227
 Qy 228 ---KQLLHNHOATNMHAIILTG--ALVTGVDKOSKSAKNIWQAVYDMEETMPLNHIVDG 281
 Db 578 WTIBEEVRFTHWNSDMPGRKPGCVAMRTGIAAGG---LWD--VLCKDCAKFVCKHWAEG 631
 Qy 282 -----SLDEGGVAYGSTAKTSTQVYFLAQRFH-----NIN 311
 Db 632 VTHBPKPPTTPEPKCPEDWGAASSRTSLFRLYAKGHEKCTWFSRDFCRAALGDDLASIN 691
 Qy 312 NLDDN-----WL-----KOH--FW--FYATTLLPGFORTYGIADSNNYWFYGPBSQLVFLD 358
 Db 692 NKCEPQQTIWRLTASGSYTHKLFWLGLTYGSPSEFTWSDGPSPVSYENWAYGEPNVYQNV 751
 Qy 359 -----KPLLRKGAGNWLAQQTTRKH-----PKDGPVNPVSTAORWSTLH 396
 Db 752 YCGBLKGDPTMSNDINCEHLMNTMIC-QIQKGQTPKPBPPTPAQDNP--PVTEDGW-VLY 807
 Qy 397 TEIWYDPLQTPPPADYGTAKHTFPWVUTYAGLIPNTQNTFVFSFKSGKLGGRAYV 456
 Db 808 KDVYY ---FSKEKETMDNARAFCKRNFGDL --VSQSESEKFKLWKVNRNDQASAY 860
 Qy 457 DIVHF---QPYSMIDG-----WRSFNPQGHEPHPDONSFTFAPNG-----QVF 494
 Db 861 FIGILLISLDKKFAMDGSKVYDYSWATGPBNFANEDECNTVMNSGFWNDINGYPNAP 920
 Qy 495 VSEALYGPKLISHN--INVLFAPSPSSQONKWPQEGCAQWLKWTGEB----- 542
 Db 921 ICQR--HNNSINATMPTMSVPSGCCKEGNFYSNKCFCFKFGMFEBERKQWQEARKA 976
 Qy 543 -VGDAAAGETITASQHGMVYFVSGEAVSSAAMRKSYRALLNSQTLVVDHIERQE 601
 Db 977 CIGFG-----GNLVSIQNEKEQAFLYHMKDSTSFTSAWTGLND----- 1013

| | | | |
|--|---|----|---|
| Qy | 602 D\$PINSVSAFHNLIDDFK\$YIPIPKFM\$R-YNG---AMNDVWDAYH\$YMF-----WF 647 | Db | 1260 -----BLLGRKG-KIVEEQTEALR-----AKIAEDKV-----NDKITE 1293 |
| Db | 1014 ---W\$N\$EHTP---LWT\$D\$RG\$H\$P\$---\$M\$A\$T\$O\$E\$A\$Q\$A\$E\$F\$K\$R\$W\$T\$Q\$V\$N\$T\$G\$Y\$P\$G\$R\$R\$S\$S\$Y\$E\$A\$D\$C\$V\$U\$G\$A\$S\$N\$E\$A\$K\$W\$M\$ 1067 | Qy | 785 AVGLILCISLVLTV\$P\$Q\$F\$Y\$V\$LS\$F\$K\$M\$R\$N\$T\$V\$A\$W\$F\$E\$L\$Y\$W\$S\$T\$-----CSQPICA 839 |
| Qy | 648 DH\$GNS-----PM\$A\$T\$O\$E\$A\$Q\$A\$E\$F\$K\$R\$W\$T\$Q\$V\$N\$T\$G\$Y\$P\$G\$R\$R\$S\$S\$Y\$E\$A\$D\$C\$V\$U\$G\$A\$S\$N\$E\$A\$K\$W\$M\$ 1067 | Db | 1294 AIA-----QWN----- |
| Db | 1068 D\$D\$TC\$DK\$K\$Y\$IC\$T\$R\$D\$S\$L\$T\$N\$P\$A\$T\$Q\$-T\$G\$F\$K\$Y\$G\$K\$S\$X\$S\$U\$M\$R\$Q\$K\$F\$Q\$W\$H\$A\$T\$Y\$C\$K\$L\$H 1126 | Qy | 840 KWTRTEAGSK\$K\$S\$K\$S\$K\$S\$K\$S\$K\$S\$K\$S\$H\$-MDLPDV\$V\$T\$S\$L\$P\$G\$S\$A\$E\$L\$K\$Q\$O\$F\$N\$S\$D\$P\$LY\$V\$R\$P\$T\$Y\$ 898 |
| Qy | 687 ---R\$IA\$V\$Y\$G\$P\$Y\$IN\$S\$S\$C\$F\$D\$S\$N\$P\$G\$Q\$O\$S\$S\$Y\$V\$N\$A\$N\$T\$E\$H\$-----W\$S\$V\$T\$D\$Y\$H\$N\$ 733 | Db | 1320 FQT\$T\$E\$L\$S\$Q\$F\$E\$M\$Y\$K\$A\$F\$E\$L\$D\$P\$S\$A\$E\$S\$S\$L\$P\$A\$-----T\$E\$E\$-----Q\$D\$P\$M\$S\$W\$W\$A\$A\$S\$T\$ 1371 |
| Db | 1127 NS\$J\$A\$S\$T\$L\$-D\$Y\$S\$N\$A\$F\$A\$M\$O\$M\$E\$T\$S\$N\$E\$R\$V\$W\$A\$N\$S\$N\$T\$D\$N\$Q\$T\$W\$T\$D\$K\$W\$R\$V\$R\$T\$N\$W\$A\$D\$E\$P\$K\$ 1185 | Qy | 899 DIP\$E\$T\$E\$L\$D\$F\$D\$A\$C\$E\$W\$K\$-V\$S\$D\$IR\$-----SG\$F\$R\$R\$G\$W\$L\$Q\$S\$V\$Y\$Q\$D\$T\$K\$U\$H\$L\$Q\$N\$T\$H\$H\$E\$P\$ 952 |
| Qy | 734 L\$K\$T\$R\$F\$Y\$N\$Y\$G\$F\$G\$P\$A\$S\$V\$A\$ 750 | Db | 1372 -----W\$R\$S\$L\$N\$D\$D\$T\$W\$T\$S\$Q\$P\$K\$R\$Q\$S\$D\$G\$K\$M\$T\$K\$-----E\$M\$P\$ 1406 |
| Db | 1186 L\$K\$S\$A\$C\$V\$U\$D\$D\$G\$T\$W\$K\$T\$A\$ 1202 | Qy | 953 NRG\$K\$L\$A\$Q\$Y\$F\$A\$M\$N\$K\$D\$K\$R\$K\$F\$K\$R\$S\$E\$L\$P\$E\$Q\$R\$S\$Q\$M\$K\$G\$F\$-D\$R\$D\$A\$E\$Y\$B\$A\$L\$R\$R\$H\$U\$V\$Y\$Y\$P\$A\$R\$P\$- 1010 |
| Db | 1407 SR\$-M\$R\$Q\$Y\$A\$A\$-----F\$E\$H\$I\$Q\$V\$Y\$A\$-----R\$E\$H\$W\$L\$K\$Y\$A\$K\$K\$P\$ 1455 | Qy | 1011 -----V\$U\$S\$L\$S\$S\$-W\$T\$K\$L\$H\$F\$F\$Q\$E\$V\$G\$-----A\$M\$B\$A\$Y\$Y\$V\$D\$P\$R\$A\$W\$-I\$Y\$S\$-L\$Y\$N\$ 1053 |
| Qy | As\$5\$4\$8\$9\$-----dynein heavy chain, cytosolic - Emericella nidulans | Db | 1456 M\$P\$F\$S\$U\$Y\$S\$U\$T\$G\$D\$W\$U\$Q\$L\$A\$S\$E\$T\$V\$R\$T\$A\$V\$A\$Q\$G\$E\$M\$A\$E\$F\$L\$K\$S\$Y\$R\$E\$-T\$W\$O\$N\$Y\$S\$D\$U\$Y\$N\$ 1513 |
| C;Species: Emericella nidulans, Aspergillus nidulans | C;Accession: A\$5\$3\$4\$8\$9\$-----dynein heavy chain, cytosolic - Emericella nidulans | Qy | 1054 K\$P\$S\$L\$Y\$S\$L\$K\$N\$Y\$P\$E\$H\$A\$K\$L\$F\$K\$E\$G\$G\$K\$C\$K\$N\$G\$Y\$A\$F\$E\$Y\$E\$P\$K\$R\$K\$E\$S\$------K\$S\$K\$N\$A\$V\$A\$ 1107 |
| R\$X\$iang, X.; Beckwith, S.M.; Morris, N.R. | Proc. Natl. Acad. Sci. U.S.A. 91, 2100-2104, 1994 | Db | 1514 Y\$O\$N\$K\$C\$K\$B\$R\$I\$R\$G\$F\$D\$D\$F\$D\$F\$A\$C\$S\$E\$-----N\$U\$N\$S\$Q\$A\$M\$R\$H\$S\$Y\$Y\$K\$E\$F\$E\$D\$A\$S\$W\$E\$D\$K\$U\$N\$R\$V\$H\$ 1565 |
| A;Title: Cyttoplasmic dynein is involved in nuclear migration in <i>Aspergillus nidulans</i> . | A;Reference number: A\$5\$3\$4\$8\$9\$; MUID: 94161539; PMID: 8134356 | Qy | 1108 L\$L\$S\$H\$W\$U\$A\$N\$T\$A\$A\$A\$R\$1\$R\$T\$D\$U\$P\$T\$S\$Y\$V\$K\$F\$------D\$V\$H\$F\$P\$O\$K\$T\$E\$T\$I\$F\$A\$F\$G\$P\$L\$S\$ 1160 |
| A;Accession: A\$5\$3\$4\$8\$9\$-----dynein heavy chain, cytosolic - Emericella nidulans | A;Cross-references: UNIPROT: P\$4\$5\$4\$4\$; GB: U\$0\$3\$9\$0\$4\$; NID: 9451538; PIDN: AAA\$1\$8\$3\$3\$1\$; PID: g\$4\$5\$1\$5\$3\$9 | Db | 1566 V\$L\$F\$D\$W\$-----L\$P\$E\$-----L\$P\$E\$ 1598 |
| A;Molecule type: DNA | C;Keywords: ATP; heterotetramer; hydrolase; microtubule binding; nucleotide binding; P-1 | Qy | 1161 PA\$--\$S\$N\$Q\$O\$L\$F\$A\$T\$S\$T\$N\$F\$Y\$P\$Y\$E\$G\$E\$S\$P\$T\$N\$T\$V\$K\$Q\$N\$L\$P\$R\$D\$E\$K\$I\$E\$N\$C\$W\$T\$-----M 1212 |
| A;Residues: I-4344 <X\$IA> | C;Superfamily: dynein heavy chain, cytosolic | Db | 1599 S\$S\$R\$F\$O\$N\$S\$B\$E\$F\$A\$M\$K\$K\$V\$K\$S\$P\$F\$V\$U\$D\$V\$A\$N\$G\$-V\$Q\$S\$E\$L\$-L\$A\$B\$U\$N\$K\$Q\$A\$G\$E\$Y\$L\$E\$R\$ 1655 |
| A;Accession: A\$5\$3\$4\$8\$9\$-----dynein heavy chain, cytosolic - Emericella nidulans | C;Keywords: ATP; heterotetramer; hydrolase; microtubule binding; nucleotide binding; P-1 | Qy | 1213 D\$R\$G\$Y\$K\$F\$ 1220 |
| F:1933-1940/Region: nucleotide-binding motif A (P-loop) | F:2228-2235/Region: nucleotide-binding motif A (P-loop) | Db | 1656 E\$R\$S\$F\$P\$R\$F\$ 1663 |
| F:2542-2559/Region: nucleotide-binding motif A (P-loop) | F:2932-2939/Region: nucleotide-binding motif A (P-loop) | Qy | RESULT 7 |
| F:1939/Binding site: ATP (Iys) #status predicted | F:2234/Binding site: ATP (Iys) #status predicted | Qy | Galactosyltransferase-related protein - <i>Thermotoga maritima</i> (strain MSB8) |
| F:2938/Binding site: ATP (Iys) #status predicted | F:2938/Binding site: ATP (Iys) #status predicted | Qy | C;Species: <i>Thermotoga maritima</i> |
| Qy | 440 NTF\$V\$F\$K\$G\$-----KL\$G\$R\$A\$Y\$D\$-----IV\$H\$F\$Q\$-----Y\$W\$ 466 | Db | C;Accession: B\$7\$2\$3\$3\$8 |
| Db | 965 NS\$Q\$E\$S\$Q\$C\$Q\$P\$K\$Y\$N\$G\$D\$T\$D\$T\$A\$Y\$N\$T\$E\$F\$P\$G\$L\$T\$E\$I\$S\$M\$R\$Q\$Y\$H\$LD\$D\$P\$Y\$O\$A\$R\$A\$T\$W\$F\$H\$ 1024 | Qy | R; Nelson, K.B.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, C.M.; Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Phillips, C.A.; Richardson, D.; C.M. |
| Qy | 467 ID\$W\$R\$S\$F\$N\$P\$G\$H\$E\$H\$P\$D\$Q\$N\$S\$F\$T\$P\$A\$N\$G\$-V\$F\$S\$E\$A\$Y\$G\$P\$K\$L\$H\$N\$N\$V\$Y\$F\$A\$P\$S\$S\$Q\$C\$N\$K\$P\$W\$ 525 | Db | Nature 399, 323-329, 1999 |
| Db | 1025 F\$D\$W\$M\$G\$T\$G\$C\$N\$L\$E\$K\$K\$S\$R\$Y\$Q\$O\$S\$T\$E\$V\$K\$Y\$C\$F\$A\$D\$I\$P\$Q\$H\$C\$T\$N\$E\$Y\$V\$A\$-----1077 | Qy | A; Reference number: A72200; PMID: 99287316; PMID: 1036071 |
| Qy | 526 EG\$Q\$G\$E\$C\$A\$O\$W\$L\$-K\$N\$-----T\$G\$E\$Y\$G\$D\$A\$G\$E\$B\$-I\$T\$A\$S\$Q\$G\$E\$M\$V\$F\$S\$G\$E\$A\$V\$S\$A\$Y\$S\$A\$M\$ 575 | Db | A; Accession: B72338 |
| Db | 1078 E\$T\$R\$K\$E\$V\$S\$Y\$D\$K\$W\$L\$Q\$F\$S\$W\$D\$Q\$S\$E\$Q\$Y\$D\$G\$D\$D\$-----1113 | Qy | A; Status: preliminary |
| Qy | 576 L\$K\$S\$Y\$R\$A\$L\$L\$N\$S\$T\$O\$T\$Y\$D\$H\$T\$E\$R\$Q\$E\$D\$-P\$I\$N\$S\$V\$A\$F\$H\$N\$D\$D\$D\$Y\$K\$Y\$P\$K\$M\$N\$Y\$R\$G\$A\$M\$ 634 | Db | A; Molecule type: DNA |
| Db | 1114 -----S\$C\$W\$L\$Q\$I\$Q\$E\$R\$K\$S\$R\$A\$T\$H\$E\$R\$Q\$S\$G\$N\$K\$Y\$D\$Y\$Q\$V\$T\$R\$Y\$N\$A\$Y\$-----1157 | Qy | A; Residues 1-104 <ARN> |
| Qy | 635 MD\$W\$-----D\$A\$H\$Y\$K\$M\$F\$W\$D\$H\$-----H\$G\$N\$S\$P\$A\$S\$Q\$A\$E\$Q\$A\$ 664 | Db | A; Cross-references: UNIPROT: Q\$9\$W\$Z\$L\$S\$; GB: A\$E\$0\$0\$1\$7\$4\$5\$; PID: 94981278; PID: AAD3583t |
| Db | 1158 -D\$Q\$K\$H\$E\$T\$T\$K\$F\$G\$S\$K\$G\$N\$R\$M\$Y\$H\$S\$E\$L\$A\$T\$A\$R\$H\$D\$E\$G\$K\$S\$K\$D\$S\$A\$S\$T\$A\$H\$A\$F\$T\$S\$V\$O\$C\$K\$R\$K\$ 1216 | Qy | A; Experimental sources: strain MSB8 |
| Qy | 665 A\$E\$F\$K\$R\$W\$T\$Q\$V\$N\$T\$F\$Q\$M\$S\$T\$R\$E\$M\$Y\$G\$Y\$T\$N\$V\$S\$C\$R\$F\$D\$S\$N\$P\$G\$Q\$S\$Q\$C\$N\$K\$W\$ 724 | Db | C; Genetics: A; Gene: TM\$7\$5\$6 |
| Db | 1217 A\$-\$-K\$W\$E\$P\$E\$D\$L\$F\$R\$Q\$G\$A\$T\$V\$R\$Q\$Y\$Q\$P\$S\$D\$M\$L\$H\$V\$E\$N\$-----1259 | Qy | A; Best Local Similarity 21.5%; Pred. No. 3.1; Mismatches 90; Indels 192; Gaps 37; |
| Qy | 725 V\$S\$T\$D\$Y\$H\$N\$L\$K\$T\$R\$F\$Y\$N\$Y\$G\$G\$F\$A\$S\$A\$D\$Q\$Q\$O\$T\$R\$F\$G\$L\$G\$T\$Q\$A\$T\$Y\$K\$P\$V\$R\$H\$D\$R\$I\$F\$P\$G\$P\$K\$F\$N\$ 784 | Db | 535 W\$K\$W\$T\$G\$B\$E\$V\$G\$D\$A\$G\$E\$T\$T\$S\$O\$H\$G\$M\$Y\$V\$-S\$G\$E\$A\$V\$S\$A\$S\$M\$R\$L\$K\$S\$V\$Y\$R\$A\$L\$L\$N\$S\$T\$ 591 |
| Qy | 941 H\$Y\$K\$M\$F\$W\$D\$H\$G\$N\$S\$P\$M\$A\$S\$Q\$A\$E\$Q\$A\$E\$F\$K\$R\$W\$T\$Q\$Y\$N\$Y\$F\$G\$Y\$Y\$V\$N\$V\$ 700 | Qy | Db 194 W\$E\$W\$A\$N\$M\$A\$V\$E\$V\$T\$N\$K\$L\$-----P\$K\$D\$G\$K\$K\$U\$R\$C\$R\$H\$Y\$E\$A\$N\$Y\$P\$E\$K\$M\$N\$Y\$V\$N\$ 242 |

| | | | | | |
|----|------|---|----|-------|---|
| Db | 295 | H1-----NFIKKDPAWMLQIGMRLKIDERTYTHIAGDQ-----BIRYANTYFKH-----338 | Db | 519 | QFFPELALCPVELEQK---YLNSALAAAEFRGLSDYEHSYPPYPPVHGRFPLND 574 |
| Qy | 701 | SCRFIDSSNGLQISLNVNN-TERIVSUTTYDHNLKTRENYLGRG-GRASVAQOGQLTRF 758 | Qy | 129 | E-TYGNNUPLLA-LYCLLCPDEKDVAFETVLEYEDRMYGKYDKWIVENAREDEVPIGHSL-- 184 |
| Db | 339 | ---FL---F1---KODAGLEKVNKLGYWEDVNADFLDKYNTLSSTHESGYNALABAMRG-----388 | Db | 575 | RFTIENILPRLSOLQKQAVEQDYGKTUVIKAIGLGSREIILKFVPTYLEGQIKVSTYL 634 |
| Qy | 759 | GLGQAIKVKEVRHD---RIIFPGPKENIAVGHLICSLVLTQFWRFY-----804 | Qy | 185 | -TGFATADFDFLYNLDDHRQKYLEKIWVITEEMEYSKVRSGKQQLHH-NHQATNMIAL 242 |
| Db | 389 | ---TKEFHINHYAGAKTQWPDLVNF-----IDEVIRLVTSDYNSCEYRSFVE 434 | Db | 635 | RVGMILLALKPLAQSKDRTYRSA-LFSILMNTAEEYE---VRVAAMILFLNEPITDMILRV 690 |
| Qy | 805 | --LSFRLKLMRWWILILVIALWFIELLDWNTCSOPTIACWTRTEALEGSKSLSSGGHNDL 862 | Qy | 243 | LTGALVLTGVDGSKKANIKQAVVYDVMEMKFLIN----HVDCG-----SLDEGVA 288 |
| Db | 435 | KNCLEKQITSIISLTV-----QLQDNMRTKNSIKHKST-TDTENSFAKIKYESKID- 486 | Db | 691 | M---AQLTNDD---PSIHYRAVLKSSIEATALEGPKWHLAKAQSYKVELVISEDFGYA 744 |
| Qy | 863 | PDVVITSIPPGSG---ASTLKKOLF-FNESDEFIYTRVPTAY-----IDPE 902 | Qy | 289 | YG-----SYATKSYTOVFLAQRHFPTINNLDNWLKQHF---W---FYYATULPGFQR 335 |
| Db | 487 | SFTIMNDLPGSKLSEFVSFLNKRARILEVGTGTGAFSIELAREADYTGIDIDP 546 | Db | 745 | YSAAWFAKKNFDSEBVTRY---RLANIGSDSLLPNLFLSTWNRPYGR---GAEN 795 |
| Qy | 903 | TELEI-----DSFVDAECWKVSDIRSFGHFLR-----GMWLSLVQDPTKLH 943 | Qy | 336 | TVGJADSNWYWFGPESQVLVEDLKPKLNKGAGWLAQQIRK-HRPKDGPMPV-STAQORMS 393 |
| Db | 547 | TSIELAIRISKDYNVENTEFVKGD---GFKLUTESSFKPQBFDLAFNMGTVVEHFKDDDLIK 602 | Db | 796 | TVGFTTSVNYQDIDPYKQILQIOPR---KSTTEHOFSADKIAKLNIKQDQLOPTESSIYK 853 |
| Qy | 944 | LQNITHLHEPRNGKLAQYFANKMDKRRKPKRRESLPEQRSEQMKGA---FDRDDEYIRALR 999 | Qy | 394 | TLHTEBYIWYDPQLTPQPPADYGTAKIHTFPNPMGIV-TYQAGLPNTQTNTFVFSKSGKL-- 450 |
| Db | 603 | MLK-----QMGEVAKPVVVGVPYSGSFYKTA-KETIAQKLGAWEYGFPERDFLTLEFLI 654 | Db | 854 | IDFKEFW-----LDQPNKIQVQSVVIFP 904 |
| Qy | 1000 | RHLVYYPSSAPVLSUSSGSTMSTLKHFFQVLTGASMRLAYTVDRAWYLSMLNSKPSLW 1059 | Qy | 451 | --GGRAYD-----IVHFQPYEWIDGWRSPFNQGHEHPBQN---SFFAPNQCVFVB 497 |
| Db | 655 | RRAGLIP-----LH---EVIGVLAEPFURRNPEW-----684 | Db | 905 | IATGBPFVFEYKEPVWVILQ---TKLSGKINYPPTKDISNNIQTEBIQFTFARN---ME 957 |
| Qy | 1060 | SLKVNPEHLAKLFKIE---GGKGKCNLNS-GYAFYEYPL--RKELSKSKSNAVSL 1110 | Qy | 498 | ALYGPKLSHLNNVLVEPAPSSQCNCNPWB---GQLGBCAQOWLKWTVGEEVGDAGEITTA 553 |
| Db | 685 | PLK-TAENLQKYFOGSEKVGSWLCPATKMGYADEFLKUDDHKKI-KFSSSTQISLIT 739 | Qy | 958 | GSVGFIDINTNQNLANIGYVVKYQINLPKINVIGQSGKFQVKL-----MMWDWDAHLYKMFNWDDHHGNSPMASI 658 |
| Qy | | | Db | 1061 | -----FCKYHGFSSSNTYRNAGDFVQDALDFSDFMSLNDIAETHVE---PTYSGK--- 1106 |
| Db | | | Qy | 554 | SQHGBMVMVSGEAVSAYSSAMRLKSVTALLNSQTLUVVDHIERQEDSPINSVSAFPH 613 |
| Qy | | | Db | 10006 | DQDLILHHSWVWPSAYQONTNTTTPSHLKTKLIRKKAEMKKELGQQQFNLL-----1060 |
| Db | | | Qy | 614 | NLDDDFKTYIPKEMNRVNGA-----MMDWDAHLYKMFNWDDHHGNSPMASI 658 |
| Qy | | | Db | | -----FCKYHGFSSSNTYRNAGDFVQDALDFSDFMSLNDIAETHVE---PTYSGK--- 1106 |
| Qy | | | Qy | 659 | QEAEQAEEBKFRWT-----QFVNVT---FOMESETITRFLAYVFYGPYINV 699 |
| Db | | | Db | 1107 | QSKTKAVTETAAVDSHNLNSTGKFGEARHRSDFVTPNSAPRRETFVKRVS---SGIKT 1161 |
| Qy | | | Qy | 700 | SSCRFD---SSNPFLQISIUNVNNTTEHVVISIVTDYHNLKTRFN-YLGFGGPASVADQGQIT 756 |
| Db | | | Db | 1162 | ATARVIDFSASFESLQ-KLEYAVTAAVAGSMVD---LKTQFAPVFMG---SQSDNGQI- 1211 |
| Qy | | | Qy | | -----RJLRGWQSLVYOPTKLNH---IHLHGPNGKLAQYFAMNKDK 967 |
| Db | | | Db | 1212 | -----NAVFKLQPKQMAPDFHKALNSAVKYLHEAD---VTFGENSNINFKGHTERSQE 1262 |
| Qy | | | Qy | 813 | WILILVIALWFIELDWISTCSCQICAKWTRAEAGSKSLS---EGHHMDLPDVVI 867 |
| Db | | | Db | 1263 | YAEQKQNSLWA-----NOCAQ-----ENAQQNEQQLGCHNVLTHSHAPDRFKASI 1307 |
| Qy | | | Qy | 868 | T---SLPGSGAEILKQFFF---SSDFLYTRVPTAVIDIPPELEID-SFVD---ACEWK 917 |
| Db | | | Db | 1308 | TYRHIAPAHTALLDSYIQLGLWSKGPEY-NPSKRLPVGQIELEANAASYDQTANVATW 1365 |
| Qy | | | Qy | 918 | VSDFRSRGIF-----RJLRGWQSLVYOPTKLNH---IHLHGPNGKLAQYFAMNKDK 967 |
| Qy | | | Db | 1366 | NGQVRFLNPLNTYITPAQIGEDSWTHPANSYSYHQEP-----PCTVDTGK 1418 |
| Qy | | | Qy | 968 | KRKFKRE---SLPE-----QRSOMK-----GAFDRDAEYIR 996 |
| Qy | 36 | VSNYSEWAVFTDDIDQ-----FKTQKVQDFRINQ---KLKXSMMLH-----PS 74 | Db | 1419 | VKETFSNRDNTLPEIWTYLMHAOTQNTWEVLVLAKRPEAKREIGKIGKLDIYISHKT 1478 |
| Db | 459 | IAKUDMWIFIRDAVTOQITOPAFQLIETWKTQKIQFLAEVISSLAKTLRHLTKDTMK 518 | Qy | 997 | ALRHHL-VVYPSARPVLSUSSGSNTLKHFFQEVTLGASMRLAVRDPRAWIYSLMLYNSK 1055 |
| Qy | 75 | LYFDPKXHADFAAKN 128 | | | |

RESULT 8
T43162
vitellogenin - gypsy moth
C;Species: Lymantria dispar (gypsy moth)
C;Date: 11-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43162 ; T43268
R;Haremath, S T ; Lehtoma, K.
Submitted to the EMBL Data Library, June 1996
A;Description: Gypsy moth vitellogenin mRNA.
A;Reference number: Z22321
A;Accession: T43162
A;Status: Preliminary: translated from GB/EMBL/DBBJ
A;Molecule type: mRNA
A;Residues: 1-1747 <HIR>
A;Cross-references: UNIPROT:Q25269; EMBL:U60186; NID:91399940; PID:gi1399941; PIDN:AA033
R;Haremath, S ; Lehtoma, K.
Arch.; Insect Biochem. Physiol. 35, 151-164, 1997
A;Title: Structure of the gypsy moth vitellogenin gene.
A;Reference number: Z223278; MUID:97468497; PMID:9327584
A;Accession: T43268
A;Status: Preliminary: translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-819, Q , 821-1747 <HIR>
A;Cross-references: EMBL:U90756; NID:g2894826; PID:gi2894827; PIDN: AAC02818.1
C;Genetics:
A;Introns: 11/1; 725/2; 796/3; 1120/1; 1410/1; 1477/2
C;Superfamily: boll weevil vitellogenin

Query Match 1.9%; Score 122; DB 2; Length 1747;
Best Local Similarity 20.4%; Pred. No. 8.2.; Mismatches 174; Indels 336; Gaps 70;
Matches 268; Conservative 174; Mismatches 536; Indels 336;

Query 36 VSNYSEWAVFTDDIDQ-----FKTQKVQDFRINQ---KLKXSMMLH-----PS 74

Query 36 VSNYSEWAVFTDDIDQ-----FKTQKVQDFRINQ---KLKXSMMLH-----PS 74

Query 459 IAKUDMWIFIRDAVTOQITOPAFQLIETWKTQKIQFLAEVISSLAKTLRHLTKDTMK 518

Query 75 LYFDPKXHADFAAKN 128

| | | | |
|--|--|---|---|
| Db | 1479 ATGKHLVNP-----YSMANNKANVKVETNAOLVADDLTY--WDDVAETPLLQVSNH 1531 | Db | 418 GGTGAOKGGSLTENABQQKPLADAVVTGQVAGEVVST-----VFESIGRNVSGENLN 471 |
| Qy | 1056 PSLSLKVNPEHLAKLFKLEGG-----KGKCNLNSGAYEPEPLAKELSKSKSN 1104 | Qy | 569 -----AYSSAAMPLKSVYRALLINSGQL-----LVDHIERQEDSPINSVSAF 611 |
| Db | 1532 PDRVLILHLSDGHLFLFDKGIFTTSQRNTTRGICQNSG----DPL---DDYKTPGL 1584 | Db | 472 LGWSFMGMFARTKAEEQIKATEK--IGLNNITMPCTNLHTFLNR-----NAF 517 |
| Qy | 1105 AVSLJSHMLANTAAALRINTDLPLTSVQLVKEBDI----VHFQKTERIFAF 1154 | Qy | 612 FHNLDIFKTYPKFNMN--RYN-GAMMDVDAHYKRNFWDEHGNSPMASTOEAEQAE-F 667 |
| Db | 1585 IVDHQSQHFGAAFT-----LDEKTNSQIQWMKIAQETATQPKLTHTVLRF 1631 | Db | 518 MYD-----IKVLPYSKETITONLGELID-----NTPSLGKVRRGIVDID 559 |
| RESULT 9 | | | |
| DB-926 hypothetical protein U0166 [imported] - Ureaplasma urealyticum | | | |
| C;Species: | Ureaplasma urealyticum | Qy | 668 KKRWTQFVNNTFQEMESTITRIAYVFGPYINVSSCRFIDSSNPGLQISLYNNTEHVVSI 727 |
| C;Date: | 18-Aug-2000 #text_change 20-Aug-2000 | Db | 560 YPGWTKVEDATGYNNSNI-----WLNTRGNFN----APCQBLEVSN----KL 601 |
| R;Glass, J.I.; Leffkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H. | Db | 562 KADLNLNDIFTNKPRTI-FSPLGLPAIAACCYHFDTDRFAKFVDDDELKVFDTELIGQRTY 660 | |
| submitted to Genbank, February 2000 | Qy | 757 RFLGLGTAQIKVPVRDRII----FPFGFK--PNIAVG-----LILCISL- 794 | |
| A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a minor | Db | 661 RDGTYVTKDKREPLYDNPSCFCAALPPPTPLSEPAVINASAKREDVNRYKVPYVTDIDIK 720 | |
| A;Accession: D82946 | Qy | 795 -----VILTFQWRFLYFSFRKLMRNLLVILYIALWFIELDWSTCSQ-----PICAKWT 842 | |
| A;Molecule type: DNA | Db | 721 GLGKDICKLTAFPSGFYSPDARLKD-----EFKQVYEAFFESNGKYTDMSQWS 768 | |
| A;Residues: 1-944 <GLA> | Qy | 843 RTEAEQGSKKSUSSEGHHMDIPDVVITSPLPGSGAETLKQLPFNSDPLYIYRPTAYDIDPE 902 | |
| A;Experimental source: serovar 3; biovar 1 | Db | 769 NYINFSTLDEINSVGTKLSPK-----PEVETDPLKNTKTNSDQGVNALQT----- 815 | |
| C;Genetics: | Qy | 903 TELEIDSPPVDAECWVYSDIRSGHFRILLGRHLQSLYQDTKLHLQNLNHLHERNRGKLAQYFA 962 | |
| A;Gene: U0166 | Db | 816 -----INETLKYGD-----NSLISATQYRKPREFYEDSSKLANLFG 854 | |
| A;Genetic code: SGCG | Qy | 963 MNKDKRKRKFKRRESLPEQRSQMKGAFDRDAEYTRALRHLYVPPSARPVLSLSSGSWTLK 1022 | |
| Matches 212; Conservative 148; Mismatches 388; Indels 461; Gaps 58; | Db | 855 Y-----GSYTNK 861 | |
| Qy | 44 VFTDIDQFKTQVQD----FRPNOKLUKKSMILHPSLYFDAGEBIQAMRQKSRAHSLHFLR 98 | Qy | 1023 LHEFFQEVLGASMRLAYIYDPRAWIYSLM-----YNSKPSLYSLKVNPFELAKLFKEG 1076 |
| Db | 9 IYDDELEGYSQRNIEDSDSLSIPBNQK----PIHLCPFR 43 | Db | 862 LNY-----LGGSNLNSDVTIAN-KVIAYSFVFKWDNFYRREWLNGLGSVNTKRNKLTFEIES 917 |
| Qy | 99 AIRSAVTMILSNPTTYLPPGRHDFAAKWNEIYGNNEPLPLAYCLLPEDK--VAPEFV 155 | Qy | 1077 G--KCKCN 1082 |
| Db | 44 ILRGFIKTTNTGECSLLVNDK--DFSTWN--IKASNEPIIAYW-----DISFTYNFKGS 94 | Db | 918 NWIQKQWTN 926 |
| Qy | 156 LEYMDRMVGYKDMLVNEAPGDEVPIGHSLTGATAFDLYNLLDNRHRQKYLEKTIWVITE 215 | RESULT 10 | |
| Db | 95 LNISK-----DVARIDPSNTLTKNPF-----KPYFQAWV----125 | Db | JC7350 |
| Qy | 216 EMYEYSKVRSGKQLLHNHOATNMIALTGA-----YTGVDKGSKKANIWQAVDVMKTMELL 275 | Db | N-Acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse |
| Db | 126 --QNYSQANNVSTIVFNDEKHTVKTITTVLKVPSVYEDIYONE-----171 | C;Species: | Mus musculus (house mouse) |
| Qy | 276 NHIVD GSL-DEGYAGSYSTAKSUTQYTF-LAQRFHNNMLDNWLRKHFWPYATLPGF 333 | C;Date: | 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 09-Jul-2004 |
| Db | 172 ---QGNLVSRAKFKQWNTDMKEYLFLPDEHAQANQ-----WLMVNNEVTFPY 217 | C;Accession: | JC7350 |
| Qy | 334 QRTYGIADENYNNWYGPESQLVLPFLKFLKNGAGNWLAQQTIRKHPKDGPMPSTAQRWS 393 | C;Comment: | This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in |
| Db | 218 STI---SKRYE---GAQITLTNVNDLTLSSGNDLILATIQCQPCDGYL-----WP 261 | C;Keywords: | glycolysis; sulfate transport; sulfotransferase; transmembrane protein |
| Qy | 394 TLHTEYIWYDPQ-----LTPOPPA-----DYGTAK1HTFPNWGVVYGA 432 | A;Reference number: | JC7350; PMID:20374462; MUID:20374462; NCBI:AB040710 |
| Db | 262 VIONQGLKTYDEREKPLGVKTSIQLLTNGPALSMLITGRTYNNNDTRELSAKRLLYKA 321 | A;Molecule type: | mRNA |
| Qy | 433 GLP-----NTQTNTEVSFKSGKLGRGRAVYDIVHQPYSIDGWRSPFNQHEDQNS 484 | A;Residues: | 1-484 <UCH> |
| Db | 322 KIPVASKFENVSAPTSNVYLYNAFFETELEYK-----DQVEKQKD 362 | A;Cross references: | UNIPROT:Q99NBO; UNIPROT:Q9EP78; |
| Qy | 485 FTAPNGOVFSRALYGPRLSH---LNNVLVYAPSSSQCNKPWE-----526 | C;Comment: | This enzyme, having a type II transmembrane topology and N-linked glycosylation sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand in |
| Db | 363 FTDPHVANIKLNGLL---LNHKWDASNNALSY--EDNRYPPXKLWDEGFKTHASFFNFKL 417 | Query Match | Score 113.5; DB 2; Length 484; |
| Qy | 527 -----GQGLGCAQWIKWTGREGEV--GDAAGEIITASQHGENMFYS--GEAVS-----568 | Best Local Similarity | 18.8%; Pred. No. 4.5; Mismatches 171; Indels 107; Gaps 18; |
| Db | 83 GSPGNLISAVEAVTQEKHIXVHATWRGSSFLGELNFNQHDPVFYLYEPWNWHL-----135 | Matches | 81; Conservative 72; Mismatches 171; Indels 107; Gaps 18; |

| | | | | | | | |
|----|------|--|------|----|------|--|------|
| Qy | 906 | EIDSFVDAECKVSDIRSGHFRILLRGWLOSLVQD---TKHLQNIHLHEPNRGLAQYFA | 962 | Qy | 279 | VGSSLDEGVAYGSYTAKSVTQVFLAQRFHENINNLDDNNWLKHFWFYYATLLEPGFQRRTVG | 338 |
| Db | 136 | -----WQA-----LYPEADESISQGAUDLFRCDFSVIRLYAOPGDPERAPPSA | 183 | Db | 482 | YDTIMK---YGYRETHCKERAKBLK-----NSQDN-----510 | |
| Qy | 963 | MVKDKKRKPKRR-----ESLPEORQSQMGAFDRDAE-----YIRALRRLHVYY | 1005 | Qy | 339 | IADSNNYMFYGPESQLYFLDKFKTFLNGAGNWLAQQIRKHRPKDGPMYPTSTAQRWSTLHTE | 398 |
| Db | 184 | -NLTTCMILFRWRNTKVICSPPLCPAAPPARADGLVEDKACECTCPCPVSLRALEAECKY | 242 | Db | 511 | -----DKESE-TFMRK-VLEPSPDDDHMPQPKKEDDINDSP-----543 | |
| Qy | 1006 | PSA-RPVULSSSSTWLKLHFFDQVL-----GASHRALYTVRDPAWIVSMLYNSKP | 1056 | Qy | 399 | YIWWYDPLQTQPQPADYGTAKIHTPPNCVVVTYAGLPNTQTN-----TPVSFKS | 447 |
| Db | 243 | PVVVKDYLRLDGL-----VLVPLLRDGLNLKVQVLERDPRAV-VANSRLKSRO | 290 | Db | 544 | -----DVEMTESDKVVRKDVEMPDIKAYAPILL--LP-TPTNDPIKADFYLRTLMSPLK | 595 |
| Qy | 1057 | SL-----YSLKNVPEHLAKLPKIEGGKGKCNLNSGYAFEEPLRKELSKSKSNASVLL | 1109 | Qy | 448 | GKLGGRAYDIVHFQPYSWIDGMRSFVPGHEHDONSFTFAPNGQVVFSEALYGPKLKSHL | 507 |
| Db | 291 | GLRESIVQVTRORGFLHVRLLAHG---VDARPGGARALPSAPPADFFUTSALEVI | 346 | Db | 596 | -----TIIHDLKVFNPL-----PPNEYTVANPKLWASYSRVFSYE-----629 | |
| Qy | 1110 | SHMLMANTAAALRINTDILP-----SYQLVKFEDIVHFQPKTTERIFAFLGIPLS | 1160 | Qy | 508 | NNVLVFPAPSQQCNKPMEQOLGECAQWLMKWTGCEBVGDAAGELEITASQHGMEMVFSGEAV | 567 |
| Db | 347 | CEAWL-----RDLLFTTRGAPAWLRLRYRLRLRYEDLWQVQAOQLRRLRFSGL-RT | 395 | Db | 630 | -EVTFP-----KDLFHECIGLKKFFDHNEXLSPS-E-TTCKHFD-ISMPSLPUV | 673 |
| Qy | 1161 | PASINQILPATSTMLFYEGE1SPTN-----TNWKWONLPRDEIKLLENICWTL | 1211 | Qy | 568 | SAYSSAMRLKSVYRALLLNSQTLVVDVHIERQ-----EDSPINSVSAPPHNLDI- | 617 |
| Db | 396 | LAALDAFAPMNTRGSSAY---GADRPFHLSARDAREAVHWWRERLSQEQRVETACAPA | 451 | Db | 674 | SATKDARELMD-YLAFFMMQMNDATPNELIEQPLPVYERMLBEDSGLLHVAQSPLTSSEIT | 732 |
| Qy | 1212 | MDRLGYPKEMD 1222 | | Qy | 618 | --DFKTYIPKFMNRYNGAMMDWDAHYKMFMWDHHGNSPMASIOEAQEAEQAFKRWTTQFV | 675 |
| Db | 452 | MRLIAYPREGD 462 | | Db | 733 | SPNFGAILRLFL---KGKUDLG-----752 | |
| Qy | | | | Qy | 676 | NVTFQMETISTR--IATVYFPGYINVSSCRFDSSNPGLQISLNVNNTTEHVVSIYTDXH | 732 |
| Db | | | | Db | 753 | NVDFNTENSNVILRLFLKLSMSVNLFPNTEVNLPHLNDLILNSLKYSTTAEEPLVY--FY | 810 |
| Qy | | | | Qy | 733 | NLKTRFNTLVGFGF---ASVADQGQTTRFGJLQTOATV/KPVHDRIIPPGFKENIAVGL | 788 |
| Db | | | | Db | 811 | LIRTLPSSIGGRFENLYRSKIPLQLVLLQSNQMLTARLPHIREY-----V | 859 |
| Qy | | | | Qy | 789 | IUCTSVLVLTFTQWRFYLSFRKLMRWILLVIAL-WFTEL---LDWNSTCSQPCAKWTR | 843 |
| Db | | | | Db | 860 | ELCITVPRVLSPVLAFLP--LNK---PLVFAUQYDPLVSGQRTLELCIDNLTAEYFD | 914 |
| Qy | | | | Qy | 844 | TEAEKSXKSLSSEGHHMDLPVVITSPLPGSGABILKQFLFNSSDFLYRVPYAYDIP-E | 902 |
| Db | | | | Db | 915 | P1PEBVTDVSKALFNLQPOPENHAISHNVRLGKLGGRNRQFL-KPPT---DLTEK | 969 |
| Qy | | | | Db | 903 | TELZIDSVDYDACEWMKVSDJSIRSGHFLRLRGWLQSLVQDYKQLHONIHHLB-----951 | |
| Db | | | | Db | 970 | TELDIDIAJD--FKINGMPEDVPLSUTPGIQSALNLTQSYKSDIHYRKSAKYLTCVL | 1026 |
| Qy | | | | Qy | 952 | --PNRGKLAQYFAMNKDK-----KRKFKEKRESL-----P | 978 |
| Db | | | | Db | 1027 | LMPKSSAEPPTNTYELKAVASIKLERIGIEKNPDLBPTVNGRDYNSOENFLRLESV | 1086 |
| Qy | | | | Qy | 979 | BQSQMKGAFDRAEYIRALRRLHVYYPSARPVLSLS--SGSWTLLKHFQEVLGASM- | 1034 |
| Db | | | | Db | 1087 | FYATSIKELKDDANDLNLNLDHFCLLQVNTTLLNKRNNGTPTNIDKNPNFMDLSLIL | 1146 |
| Qy | | | | Qy | 1035 | -----RALYI--VRDPAWIYSMLYNSKPSLV-----SLKVPHEHLAKF-----1072 | |
| Db | | | | Db | 1147 | DAIPFALSYYIPEREVGVLAYKRIYESCLIGEELALSHSF1PE-LAKQFHILCYDET | 1205 |
| Qy | | | | Qy | 1073 | --KIEGG-----KOKCNUNISGAYAPEYEPLRKELSKSKSNASVSLSH | 1111 |
| Db | | | | Db | 1206 | YNNKRGGVGIKVLIDNVKSSSYFLKQYQYNLANGLF-----VLDPTQSEAPS--- | 1254 |
| Qy | | | | Qy | 1112 | LWLNATAAALRINTDILPTSYQLYKPFEDIVHFRQKTERIFAFGLGIPLSPLASLNQILFAT | 1171 |
| Db | | | | Db | 1255 | -----AITDSAEKULIDLSSITFADVKEED-----LG-----NKVLENT | 1288 |
| Qy | | | | Qy | 1172 | STNFYLPYEGESPTPNVWKNLPRDEIKLKENICVTLMDRUGP--KFMD | 1222 |
| Db | | | | Db | 1289 | LTDIVC----ELSNaNPKV--RNACQKSLHTISNL-----TGIPIVKLMD | 1327 |
| Qy | | | | Qy | | | |
| Db | | | | Db | | | |

RESULT 11

S46715 hypothetical protein YHR099w - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: hypothetical protein H9332.1

C:Species: *Saccharomyces cerevisiae*

C:Accession: S46715 #sequence_revision 28-Oct-1994

C:Accession: S46715 #sequence_revision 28-Oct-1994

R.Vaudin, M.

submitted to the EMBL Data Library, May 1994

A:Description: The sequence of *S. cerevisiae* cosmid 9332.

A:Reference number: S46715

A:Accession: S46715

A:Molecule type: DNA

A:Residues: 1-3744 <VAL>

A:Cross-references: UNIPROT:P38811; EMBL:U00060; PID:AAB68923.1; PID:g4879

C:Genetics:

A:Gene: SGD:YRA1

A:Cross-references: SGD:S0001141; MIPS:YHR099w

A:Map position: 8R

Matches 256; Conservative 196; Mismatches 482; Indels 439; Gaps 66;

Score 113.5; DB 2; Length 3744;

Best Local Similarity 18.6%; Pred. No. 1.1e+02;

Qy 26 MFAFSTFEE---SVSNYSBEWAFTDDIDQFKTORQDFRPF-----NOKLKCSMLHPSL

Db 218 MFSPKILSBCPITMVTLISSY-----KQLTSTSPLPFTPLMNLNIQIKOCQ-----265

Qy 76 YFDGEIQNMROKSRASHLHFLAIRSAVTMVLNSNPTY---YLPPIPQKHADF-----123

Db 266 --EAR-----QAERSGEH-----FTSISTEIIINRPAYCFDIFIQIKATSFLAYFIRGY

313

Qy 124 AAKNEIYGNLNPPLALYCLL-CPEDKVAPEFVLEYMDRMV---GYK-----DWLVEN

172

Db 314 APEFQDQYVFNPVDIIRLUQDCSELSSARKELLHATHHILSNYKKLFLPKDLYD

372

Qy 173 APGDPEVPIGHSLT-----GPAFAFDPLYNLLDNHRRQKYLEKIIWTEEMYEY-----221

Db 373 --ERILIGNGFTMHBTLRPLAYSTADPHINSELQSLSEIETKITYGYLIDESL

429

Qy 222 KVRSGKQJLHHNQATNMIAALLTGALVTSWKGSKANIWKAQVYDVMKTM-----FLNHI

278

Db 430 TVQIMSAKLNU-----LVERILKLGRENPOEARPKLMIIDSYMNRFKTLNQR

481

| | | | |
|--|--|--|--|
| RESULT 12 | S 783 98 hypothetical protein 2216 - beechdrops plastid | Db | 839 NIPHRSEIIVELKDNDQI-----CNPFL-----EPDL-KIVHLKCRKPFL 881 |
| C;Species: plastid Epifagus virginiana ('beechdrops') | Qy | 561 FVSGEAVSAYSSAMPLKSKYRALLLNQSOLVVHDIER.QEDS--PINSVSAPPHNLDI 617 | |
| C;Date: 25-1998 #text_change 09-Jul-2004 | Db | 882 -----LGYCTTSRKLK-----LLITGGRFLPKPRCMIDSPHTINRSKSFDNTD 928 | |
| C;Accession: S78398; S78404 | Qy | 618 DFKYTPYKFMNRYNGAMMDVDAHYKMFWDHAGNSPMAISQEAFOQAEEFKKRWTQFVN- 676 | |
| R.; Wolfe, K.H.; Morden, C.W.; Ems, S.C.; Palmer, J.D. | Db | 929 -SYLSMIFNK----DTWLNLYKP----HRSSLISFYKAN----LQFLNN 968 | |
| J. Mol. Evol. 35, 304-317, 1992 | Qy | 677 -----VTFOMEESTTRIAVYFGPYINVSSCRFDIDSNPGLQISLAVNNTTEHV- 724 | |
| A;Title: Rapid evolution of the plastid translational apparatus in a nonphotosynthetic | Db | 969 PHNFCFYCNTRLPPVKAHHNYFTYGOFLNL FIR----NKIFSLCVDKKKHAF 1021 | |
| A;Reference number: S78398 | Qy | 725 -----VSTV-----TDYHNLTRENY----LGFFGGFASVADQ 752 | |
| A;Accession: S78378; MUID:93021155; PMID:1404416 | Db | 1022 WGGRTDTISPIESQVKIFIPKNFQPSGDETYENLSQLQPFFPSRSDPFVRLJANIYGTPLTE 1081 | |
| A;Status: nucleic acid sequence not shown; translation not shown | Qy | 753 GQITRPGLGQAIVKPVRHDIIFPG----FKPVAVGL-----789 | |
| A;Molecule type: DNA | Db | 1082 GQIVNLG--RTYCQPLSDMNLSDSEGKPNFHQYLFNFNSNMGLIPTCPDKYLPSBKRKCR 1138 | |
| A;Cross-references: UNIPROT:P30072; EMBL:M81884; NID:9336917; PID:9336938 | Qy | 790 -LCIS-----LYTLTFLQRFLYSFKLMRWILILVIALWFI----824 | |
| A;Genetics: G1 | Db | 1139 SLCINKYKCVERGKGMYRTFORK--VAFTSLSKWNLFQTYMPWFITSAGYKVINLIFLDTF 1196 | |
| A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992 | Qy | 825 -ELDWTWSTCSOPICAKWTRTEAESKSKLSSEGHMIDL.PDVWVTSPLPSGAEILKQLFF 883 | |
| C;Genetics: <G1> | Db | 1197 SELLSLSS-----SKFVY-----IFNNHMGS-----1221 | |
| A;Gene: 2216 a | Qy | 884 NSSDFLYIRYPTAYIDIPETELEIDSFDACEKVSDIRSGHFLRLGQSLVQDTKLH 943 | |
| A;Genome: plastid | Db | 1222 -----TSWRLINKKRCLPQPNL-----1SEISSKCLHNLLSETRQNNEESP 1264 | |
| C;Genetics: <G2> | Qy | 944 LQNHHLHEPN-----RGKLAGYFAMNKDKRKFKRRESLPPQRSQM 984 | |
| A;Gene: 2216_b | Db | 1265 LISTHRSVPNREFLYSILFULLVGVLYRTHLFLVSRASSELQTEFRKVSKLMLPMSMI 1324 | |
| A;Genome: plastid | Qy | 985 KGAFDRDAEVTRALRHLYVPPSARPVLSNSGSGWTLKHFQEVLGASMRLAYLVRDPR 1044 | |
| C;Keywords: plastid | Db | 1325 B-----LKLNLNRYPT----PASNFSW-----LKNLFLIVA---1350 | |
| Query Match | Qy | 1045 AWIYSMLYNSKPSLVSLSLKNVPEHLAKLF-KIEGGKGKCNLNSGYAFYEYBL-----1094 | |
| Best Local Similarity | Db | 1351 -----MEQLVYSLSEEETRASGENLGPAYGKSIISRNKYNVNLIDLIPNIRIL 1401 | |
| Matches | Qy | 1095 -----BKELSKSKSNAVSLLSHL----WIAINTAAALRINTDLOPTSYQLVKFBDIVHFP 1144 | |
| 234; Conservative | Db | 1402 PSRNMRRHLSHTSKETYSLSLRKRKNVNGDWIDIESWANSDDER--EFLVQFS 1457 | |
| 187; Mismatches | Qy | 1145 QKTTE-RIF AFL 1155 | |
| 42; Insertions | Db | 1458 ALTTEKRYQIL 1469 | |
| 408; YSTVSNEDESEYHTLINV-----EFV-----LEYMDRMVGK-----DWLVENAPGDEVIGHSL 184 | Qy | 226 WGKOLLNHNQATNMIALLTGALVTGVDGSKANIKWAVDVMWEK-----MFLL 275 | |
| 466; EHEKQMINTHMLPPEEEFGNPTRLV-----EFKPLKSFIDPSFLQETATEESDQVQKRSYSSTLFT 465 | Db | 640 FS----IYMDQKRNVVFASNNNMEAQNQR---FIRNMKIQSYIYGYIIRNVLHREFLM 692 | |
| 147; EDKVAF-----EFV-----LEYMDRMVGK-----DWLVENAPGDEVIGHSL 184 | Qy | 276 NH1VDSGLSDEGVY-----GSYTAKSVTVQYFLAORHF-NINNUDDNNMLKMHFWFYYA 327 | |
| 522; QDLSFVPSRSRENKELNWKILKITYLKVNTVSHPISSDPGDGVLKDEP--DMDSNNKI 579 | Db | 693 NR-----SDHNLNEYEIKRDQIGDTLNRTIKYMINQHLSNPKKCONKWNPILF----743 | |
| 185; TGF-ATAFDPFLYILDNHR-----OKYLEKIVW----ITEMYEYSKVRS 225 | Qy | 328 TLJPGFORTVGIADSNTNWYGPESOLVFLDKPILKNGAGNLAQQTTRKRPKDGPMPVPS 387 | |
| 580; SVFNKNTFPLFLPHDWNRYGTLYNLHDFELEERQKADELFTLTSITEPDLYVHKGFS 639 | Db | 744 -----FSSTERSYRN-----RYKRSNSNPNLE-----771 | |
| 226; WGKOLLNHNQATNMIALLTGALVTGVDGSKANIKWAVDVMWEK-----MFLL 275 | Qy | 748 TAQRWSTLNTEYIWYDPOLTPQQPADYGTAKIHTFPNWGVVTYAGL.PNTQNTFVFSFKS 447 | |
| 640; FS----IYMDQKRNVVFASNNNMEAQNQR---FIRNMKIQSYIYGYIIRNVLHREFLM 692 | Db | 772 -----HLEH-----FVSEOK-781 | |
| 276 NH1VDSGLSDEGVY-----GSYTAKSVTVQYFLAORHF-NINNUDDNNMLKMHFWFYYA 327 | Qy | 448 GRKUGRAYDIDVIFQPYSWIDGPRSFSNFGHEAHDDQNSSTFAPNGQYVSEAL-----YG 501 | |
| 693 NR-----SDHNLNEYEIKRDQIGDTLNRTIKYMINQHLSNPKKCONKWNPILF----743 | Db | 782 SHFKFKVIFPLDIRENOYS-ID-WSAFIDTKDLISKPLRF-FLSKULLPFLNSLPPFCVSFG 838 | |
| 328 TLJPGFORTVGIADSNTNWYGPESOLVFLDKPILKNGAGNLAQQTTRKRPKDGPMPVPS 387 | Qy | 502 PKLSHLNNVLF-APSPSSQCNKWPWEGOLGECAQWLMKWTGEEVGDAAGEITASQHGEWV 560 | |

RESULT 13
S76412
hypothetical protein s1r0408 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S76412
R.; Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Miyajima, N.; O., K.; Okamura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, O.; Reference number: S74322; MUID:97061201; PMID:8905231
B;Accession: S76412
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-4199 <KAN>

A; Cross-references: UNIPROT:P74440; EMBL:D90915; GB:AB001339; NID:91653604; PIDN:BAA1854
 A; Note: the nucleotide sequence was submitted to the BMBL Data Library, June 1996
 C; Superfamily: Synechocystis hypothetical protein B1r0408

Query Match 1.7%; Score 112.5; DB 2; Length 4199;
 Best Local Similarity 20.9%; Pred. No. 1.5e+02;
 Matches 94; Conservative 42; Mismatches 131; Indels 183; Gaps 25;

Qy 224 RSWGKOLLH-NHOATNMIALLTGALVTGVDGSKANI-----WKOAV 264
 Db 3062 QSWGLPAAEINELAVNNNTSGNPQIAGL-GKGKIEYYNGSTWVNGPYQGDGWRSAI 3119

Qy 265 VDVMETKTMFLNHTVGSDE--GVAYGSYT-----AKSVTQYVFLAQRH 307
 Db 3120 -----TOMVQWGDGPSPQIVYGLADGAVIYNTTOSGRHTINNGFKGSVTLQVQWQEA 3173

Qy 308 FNIN--NLDDN-----WIKMH--FWPYATLLEFGPT-----VGIADS 342
 Db 3174 SNPNTVUGLDNSEVQYYQGSNGWNTQHDGWWYFVQQL-AVQTSNDAAQPLVYVNGLBD 3232

Qy 343 NYN---WFY---GPESOLVFLDKFLKQAGNVLAAQOIRKHPRKDGMVPSAQRWSTL 395
 Db 3233 NGNNNSVWYVQGSQEGGWTFSLG-LPSAA-----IAQMVAQWNF- 3272

Qy 396 HTEIYWDPPDQLTPQRPADYGTAKIHTEPNNGVWYTAGLPLNTQNTFVFSKSGKLGGRAV 455
 Db 3273 -----SSSPNPNVNDLKI-----VV-----GQdStvSYNG-----3301

Qy 456 YDIVHFPQYSWIDGVR---SFNPQHBEHPDONSFT--FAPNGQVFVSEALYGPKUSHLNV 510
 Db 3302 -----NGMTATPPAINSSLQIPTNAITQWSANGQPOQTIVGLGDPEDYD-NCQ 3347

Qy 511 LVFAPSPPS-----SQCNKEWGEOLGECAQ-----WLKWT 539
 Db 3348 LWLPNPSQSWQBLOGSVNATSPITOQDSSWTESLVPNSSQTDNLISYVFFGSDFNDTVNQF 3407

Qy 540 GE----EVGDAAGEGITIASQHGENMFVSG 564
 Db 3408 GTIGDDVMVSATGGSFLAGGGDDDLTG 3437

RESULT 14 AE1398 iso-tRNA synthetase [Imported] - Listeria innocua (strain Clip11262)
 C; Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C; Accession: AB1698
 R; Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baguero, F.; Berche, P.; Biceckel, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dusburgert, O.; Entian, K.D.; Fsihi, H.; Jones, J.M.; Karst, U.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative Genomics of Listeria species.
 A; Reference number: AB1077; MUID:21537219; PMID:11679969
 A; Status: Preliminary
 A; Molecule type: DNA
 A; Cross-references: UNIPROT:Q929Z6; GB:AL592022; PIDN:CAC97357.1; PID:g1614641; GSPDB:G
 A; Experimental source: strain Clip11262
 C; Genetics:
 A; Gene: ileS
 C; Superfamily: isoleucine-tRNA ligase

Query Match 1.7%; Score 111; DB 2; Length 921;
 Best Local Similarity 18.8%; Pred. No. 19;
 Matches 174; Conservative 112; Mismatches 306; Indels 334; Gaps 46;

Qy 117 PPKHDAFAAKRNNEIYGNLPLIALCYCLCPEDKVAEEFYLEMMDRMVGYDWLIVENAP- 174
 Db 21 ENKEPEWQAKW-----EEKELYKEKIQB---KNAGRRAYILHDGPPY 58

RESULT 15 C89874 autolysin [Imported] - Staphylococcus aureus (strain N315)
 C; Species: Staphylococcus aureus
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C; Accession: C89874
 R; Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Uji, Y.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001
 A; Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
 A; Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: C89874
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1248 <KDR>
 A;Cross-references: UNIPROT:Q99Y41; GB:BA000018; PID:913700854; PID:BA842150.1; GSPPB:C
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: atl

Query Match 1.7%; Score 110.5; DB 2; Length 1248;
 Best Local Similarity 18.3%; Pred. No. 32; Mismatches 66;
 Matches 245; Conservative 146; Indels 487; Gaps 66;

Qy 30 STPFEVSXNSXSEWAFTDDIDQPFXTQ---KVQDFRPNOKLKSMLHP--SLYFDA---- 79
 Db 187 TTFASQAQPRSVAAFTPKTSLPKTPQVNMISSINDYIRONNLKAPKIEEDDTSFPKYAYRN 246
 Qy 80 --- -GETOAMRQKSRSRASHLHLF---RAIRSAAVTYMLS---- 109
 Db 247 GVRPEGITVHDTANDRSTINGELSYMKNYQNQAFHAFVGDRIETAPPDLSMGVGA 306
 Qy 110 -NPTTYLPPKH-----ADFAAKRNNEYGNUNPLPLAXCLLPEDRYAFE 153
 Db 307 VGNPFRINVETIVTHDYASFARSMMNNYADYATQLOIYG---LKPD\$---- 350
 Qy 154 FVLEDYMDRMVGKDWLIVENAPG--DEV--PIGHSLTGATFADFLYNLLDNHRQKYLEKI 210
 Db 351 -ARYDGNGCTWTHAVSKYLGTDHADEHGV--LRSANYSHDOLYLIN---EKLIK 403
 Qy 211 WVITEEMYEYSKVRSGWKOLLRRH---- 247
 Db 404 -----GKVAPWGTGTTTPTPSKPTTBSKPTSGKLTVAANNGVAIKPTNSGLY 453
 Qy 248 VTGYDKGSKANIINKOAVNDMEXTMELLNHTDGSILDEGVAGSYTAKSVTOVFLAQH 307
 Db 454 TTIVDQTKGHA-----TNEVQKTEAV---
 Qy 308 FNINNLDNNRKGKHFWMFYATLIP-----
 Db 491 NSGNKP--GWWKEGDDVVNTAKSPVNTNQSYSTSFKSGTFLYTPWGISKQAVGSVSGGNQ 548
 Qy 346 WFGCPESOLVFLDKFIL---KNGAGAWNL--AQOIQKRHRPKDGPW-VPSTAQRWSTLHTE 398
 Db 549 TFRASKQQQI--DKSIYLYGSVNGKSGHNSKAYLVDTAKTPTPKPESTPTNNKLTVS 606
 Qy 399 YDVIDPQLTPQPADYGPFAKIRFPN--WGVUTYGAFLPNQT-NTEVFSRSGKLGGRAY 455
 Db 607 SL-----NGVAQINAQNNGHFTTYDKGKPTKEVQKTPAVTKASLGCGNKF 653
 Qy 456 YDIVHFQPYSWIDGWRSPNPGHBPQNSFTFAPNGQVFSEALYGPKLSHUNLVFAP 515
 Db 654 YLV-----KDYNSPTLI--GNYKQGDVYNNAKSPVNVMQTYV 690
 Qy 516 SPSSQC-NKPWEQOLGEAQWLKMTGEEVGDAEGEITIASQHGMVFV-----SCEAV 567
 Db 691 KPGTKLYSYFW-GTYQEAGAVSGTGNQFRATKQ---QIDKSIYLFGTVNGKSGWVS 745
 Qy 568 SAY-----SSAMRLISVYRALLLNSQLLNUVDHIERQEDSPINSVSAFHNLIDF 619
 Db 746 KAYLAVPAPKRVAQPKTAKVAYTUVKQPQTQTVSKIAQVK--PNNT----- 791
 Qy 620 KYIPYKFMNRYNGAMMDWDAHYKMEFWDHGNNSPMASIEAQAAEFKKRWTQFVNVTF 679
 Db 792 -----GIRASYE-----KTAKGAKYADR-----TF 813
 Qy 680 QMESTITRIAYVFYGPYINSSCRFDSSNPGQIISLNVNTEHVWSIVTVDYHNLTREN 739
 Db 814 -----YVTERAH-----GNBTYVLLNNTSH--NIPGMENVKD-LN 847
 Qy 740 YLGFCGGPASVADGQITRGFLGQTAIVKPRVHDRIIFPGFK-----FNIAVGLILCI 792
 Db 848 VQNLGKEYKTTOKYTWNKSNGLSMV-----PWGTKNQVILTGNNTAQGTFENAT 896

Search completed: June 23, 2005, 08:50:24
 Job time : 34.711 secs

Qy 793 SLVLTFOWRPFYLSPRKLMRMLILVIALWIELDWWSTCSOPICAKWTRTEAEGSSKK 852
 Db 897 KQVSUCCDVYVLYGTINNRTGMVNA-----KOLTAFTAVKPTSAAKDNYT 942
 Qy 853 L---SSEGHMHDLP--DVVITSLPQSGAEILKQLPFNNSDPLYTRYPTAYDIPETEI 907
 Db 943 YVIKNGNYYVTPNSDATTYSLKA-----FNEQPAVVK----- 977
 A;Gene: atl

Qy 908 DSFVDACEWKVSDIRSGHFRLLRGWLSQYDFTKLUHLQNTLHLPNPRGKLAQYFAMNKDK 967
 Db 978 EQVINGTWWYGGKLNSGKL---ANIKS---TDIAKELI----- 1009
 Qy 968 KRKEPKRERESLPEQRSMQKGAFDRAEYIARLRHLYVYPSARPVLSLSSGWSWTLKHFFQ 1027
 Db 1010 -KYNQUTGMUINQVAGIQAQ-----JQXPKPVQR-----PEKWTDA--NPFN 1047
 Qy 1028 EVLGASMRALTYVDRDRAFTYSMLNSKPSLNVPEHLAKLPKIEGGKGKCNL-NSG 1086
 Db 1048 DVKHA-MDTKRLAODP-ALKYQFLRDQPNONISIDKINQFL-----KGKGVLENQ 1096
 Qy 1087 YAFEYBPLRKELSKSKSNAVSLSHIWMLANTAAALRINTOLLPT---SYOLVKEFDIYHF 1143
 Db 1097 AAP-----NKQAQMYCINEYLIASH-----LLETENGTSQIAKGADVNN- 1136
 Qy 1144 POKTTERIFAFGLIPILPUSPASLNQILFATSTNLFLYLPYEGEISPTNTNWKONLPRDEIKL 1203
 Db 1137 -----NKVY--TNSUTKCYHNFQTAAYDNDPL-----REGIKY 1167
 Qy 1204 IENICWTLMDR--LGYPKFM 1221
 Db 1168 AKQAGMDTWSKAIIVGCAKFI 1187

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|-----------------------------|
| 1 | 6527 | 99.9 | 1222 | Q8IZUB homo sapien |
| 2 | 186 | 64.1 | 784 | Q6p523 homo sapien |
| 3 | 4037 | 61.8 | 755 | Q9HAK5 homo sapien |
| 4 | 2983.5 | 45.7 | 597 | Q8BZP3 mus musculus |
| 5 | 2597 | 39.8 | 480 | Q9n511 homo sapien |
| 6 | 1757.5 | 26.9 | 958 | Q8BLI4 mus musculus |
| 7 | 1753.5 | 26.8 | 958 | Q9UL01 homo sapien |
| 8 | 8 | 20.4 | 314 | Q9had7 homo sapien |
| 9 | 1242 | 19.0 | 230 | Q6j98 gorilla gorilla |
| 10 | 1194 | 18.3 | 222 | Q6J9G9 pan troglodytes |
| 11 | 215.5 | 3.9 | 680 | Q8C1W4 streptomyces |
| 12 | 192 | 2.9 | 776 | Q8UBJ1 agrobacteri |
| 13 | 150 | 2.3 | 411 | Q9EOC0 mus musculus |
| 14 | 147 | 2.3 | 411 | Q93916 homo sapien |
| 15 | 145 | 2.2 | 2504 | Q85160 photobacterium |
| 16 | 143.5 | 2.2 | 761 | Q9KWT4 sphingomonas |
| 17 | 134 | 2.1 | 483 | Q794G9 mus musculus |
| 18 | 134 | 2.1 | 530 | Q8B2T6 mus musculus |
| 19 | 134 | 2.1 | 530 | Q80WV3 mus musculus |
| 20 | 133 | 2.0 | 483 | Q9UED5 homo sapien |
| 21 | 133 | 2.0 | 530 | Q9Y4C5 homo sapien |
| 22 | 132.5 | 2.0 | 1337 | Q8A3U4 bacteroides |
| 23 | 129 | 2.0 | 304 | Q677T4 lymphocyte |
| 24 | 128.5 | 2.0 | 532 | Q6D4RS erwinia carotovora |
| 25 | 128 | 2.0 | 486 | Q9VMC3 drosophila |
| 26 | 127 | 1.9 | 441 | Q93403 torpedo californicus |
| 27 | 127 | 1.9 | 1456 | Q8NHR1 manzanares |
| 28 | 127 | 1.9 | 2958 | Q7RNZ3 plasmid |
| 29 | 126.5 | 1.9 | 420 | Q6D9Y9 brachyainfluenzae |
| 30 | 126 | 1.9 | 388 | Q9WUE5 mus musculus |
| 31 | 126 | 1.9 | 2019 | Q6BEW0 caenorhabditis |

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03:
1: uniprot_sprot;*
2: uniprot_trembl;*

OM protein - protein search, using SW model
Run on: June 23, 2005, 08:33:13 ; Search time 134.102 Seconds
(without alignments)
4666.302 Million cell updates/sec

Title: US-10-697-828-15
Perfect score: 6532
Sequence: 1 MPKGAPPWIMALMFTGHLL.....LIENICWTLMDRGYPKFMD 1222

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:

1612378

RESULT 1

QBIZUB PRELIMINARY;
ID Q8IZUB ;
AC (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DB NCAGI .

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606 ;
[1] ;
RN

RP SEQUENCE FROM N.A.

RA Goossens D., Del-Pavero J., Van Broeckhoven C. ;
Submitted (PEB-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF480435; AAN32995.1; -
DR GO: GO_0008146; F: sulfotransferase activity; IBA.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase_1.1.
SQ SEQUENCE 1222 AA; 140272 MW; 1568839784E99AE9 CRC64;

Query Match 99.9%; Score 6527; DB 2; Length 1222;
Best Local Similarity 99.9%; Prod. No. 0;
Matches 1221; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MPKGAPPIMALFTGHFLALLMAFSTFEESVSNSSEAWFTDDQFKTQVKQDF 60
Db 1 MPKGAPPIMALFTGHFLALLMAFSTFEESVSNSSEAWFTDDQFKTQVKQDF 60
Qy 61 RPNQKLKKSMLHPSLYFDAGEIQAMRQKSRASHLHFLAIRSAVYMLSNPTVYLPPPKH 120
Db 61 RPNQKLKKSMLHPSLYFDAGEIQAMRQKSRASHLHFLAIRSAVYMLSNPTVYLPPPKH 120
Qy 121 ADFAAKNEIYGNLNPPLAYCLCPDKVAFPEVLEYMDRMGYKDMLVENADGEVEPI 180
Db 121 ADFAAKNEIYGNLNPPLAYCLCPDKVAFPEVLEYMDRMGYKDMLVENADGEVEPI 180
Qy 121 ADFAAKNEIYGNLNPPLAYCLCPDKVAFPEVLEYMDRMGYKDMLVENADGEVEPI 180
Db 181 GHSLTGATAFDFTLYNLDPNRRQXLEYLWVITBEMVYYSKVRSGKQLHHNQATNM 240
Db 181 GHSLTGATAFDFTLYNLDPNRRQXLEYLWVITBEMVYYSKVRSGKQLHHNQATNM 240
Qy 241 ALLTGALVTGVDKGSKANIKWQAVDNEKTMFLINHIVDGSLDEGVSYTAKSVTQY 300
Db 241 ALLTGALVTGVDKGSKANIKWQAVDNEKTMFLINHIVDGSLDEGVSYTAKSVTQY 300
Qy 301 VFLAQRFHENINNNLDNNLKMHWYFVYATLPGFORTVGADSNYNNPKGPESOLVPLDKF 360
Db 301 VFLAQRFHENINNNLDNNLKMHWYFVYATLPGFORTVGADSNYNNPKGPESOLVPLDKF 360
Qy 361 ILRNGAGNWLAQQRIRKHPKDGPMVPSAQRWSTLHTETIYWDQLTPOPADGTAKIH 420
Db 361 ILRNGAGNWLAQQRIRKHPKDGPMVPSAQRWSTLHTETIYWDQLTPOPADGTAKIH 420

ALIGNMENTS

| Line | Sequence | Result | Score | Length |
|------|---|--------|-------|--------|
| 421 | TFPNWGVITYAGLPIQTNTFVSKSGKLGGRAYDVHFPQYSWIDGWSFNPGHEHP | 480 | | |
| 421 | TFPNWGVITYAGLPIQTNTFVSKSGKLGGRAYDVHFPQYSWIDGWSFNPGHEP | 480 | | |
| 481 | DQNSFTFAPNGQVESEALYGPKLSSLANVLVFAFPSPSSQCNKPWEGOLGECAOWLKNTG | 540 | | |
| 481 | DQNSFTFAPNGQVFSEALYGPKLSSLANVLVFAFPSPSSQCNKPWEGOLGECAOWLKWTG | 540 | | |
| 541 | EFGDAAEGLITASQHGEMIVFSGEAVSAYSSAMRLKSYTARRLLINSOTLIVDHIERQ | 600 | | |
| 541 | EFGDAAEGLITASQHGEMIVFSGEAVSAYSSAMRLKSYTARRLLINSOTLIVDHIERQ | 600 | | |
| 601 | EDSPINS/SAFFNLIDPKYIYPKEMARYNGAMMDWDAHYKNWFIDHGNSPMASQB | 660 | | |
| 601 | EDSPINS/SAFFNLIDPKYIYPKEMARYNGAMMDWDAHYKNWFIDHGNSPMASQB | 660 | | |
| 661 | AEQAEFKRKWTQFVNNTFQMEESTITRIAYVFYGPYINVSCTRDSSNPGLQTSLYNN | 720 | | |
| 661 | AEQAEFKRKWTQFVNNTFQMEESTITRIAYVFYGPYINVSCTRDSSNPGLQTSLYNN | 720 | | |
| 721 | TEHVISVITYDHNLTKRTRNLYLGFCCFAVASDQGQITREGLGTQAIKVPKVRHDRLIFPPGF | 780 | | |
| 721 | TEHVISVITYDHNLTKRTRNLYLGFCCFAVASDQGQITREGLGTQAIKVPKVRHDRLIFPPGF | 780 | | |
| 781 | KENIAVGLLICISLVILIPQWRFLISPKLMLRNLILITALWFLDWDSTCSQPICK | 840 | | |
| 781 | KENIAVGLLICISLVILIPQWRFLISPKLMLRNLILITALWFLDWDSTCSQPICK | 840 | | |
| 841 | WTRTEAEGSKKSLSSEGHMDLPDVVITSLPGSAAEILKQLFFNSSDFLYIRVTAYIDI | 900 | | |
| 841 | WTRTEAEGSKKSLSSEGHMDLPDVVITSLPGSAAEILKQLFFNSSDFLYIRVTAYIDI | 900 | | |
| 901 | PETELIEDSFVDAEKWYKSDIRSCHFRLLRGWLQSLVQDPTKHLHQNIHLHEPNRGKLAQY | 960 | | |
| 901 | PETELIEDSFVDAEKWYKSDIRSCHFRLLRGWLQSLVQDPTKHLHQNIHLHEPNRGKLAQY | 960 | | |
| 961 | FAMNKDKRKFKRESLPEORSONQKGAFDRAEYIARLRHVLYTPSARPVLSLSGSMT | 1020 | | |
| 961 | FAMNKDKRKFKRESLPEORSONQKGAFDRAEYIARLRHVLYTPSARPVLSLSGSMT | 1020 | | |
| 1021 | LKHIFQEVLGASMRALYTVDPRAWIYSMLYNKSPSLYSLKNVPHEIAKLFKIEGGK | 1080 | | |
| 1021 | LKHIFQEVLGASMRALYTVDPRAWIYSMLYNKSPSLYSLKNVPHEIAKLFKIEGGK | 1080 | | |
| 1081 | CNLSNGFAFEYPEPKRELKSKSKSNAVSLSHLWNTAAALRINTDILPLTYSQVKFEDI | 1140 | | |
| 1081 | CNLSNGFAFEYPEPKRELKSKSKSNAVSLSHLWNTAAALRINTDILPLTYSQVKFEDI | 1140 | | |
| 1141 | VHPQQKTERIIFAGLGIPLSPASLNQILPATSTNLFLPYEGEISPTNTNWQNLPRDB | 1200 | | |
| 1141 | VHPQQKTERIIFAGLGIPLSPASLNQILPATSTNLFLPYEGEISPTNTNWQNLPRDB | 1200 | | |
| 1201 | IKLJENICWTLMDRGYPKFMD 1222 | | | |
| 1201 | IKLJENICWTLMDRGYPKFMD 1222 | | | |

| | | | | | | | |
|-----------------------|--|--|--------------|------------|---------|---|-----|
| Qy | 611 | FPHNLDIDFRYIPYKFMNRNGAMMDVDAHYRMNEWFDHGNNSMASIQEAEQAAEFKKR | 670 | | 61 | HIVDGSLLDEGVAGSYTAKSVTQYVFLAQRFHNINNLDDNNWLKMHFWFYATLPGFORT | 120 |
| Db | 601 | FFNLDIDFKYIPYKFMNRNGAMMDVDAHYRMNEWFDHGNNSMASIQEAEQAAEFKKR | 660 | Qy | 337 | VG1ADSYTNWFGPESOLVFLDKFILNGAGNWLAQOTRKHRPKDGMPVSTAQRWSTLH | 396 |
| Qy | 671 | WTQFVNNTFCOMEPLITRTRIAVYFQPYINVSSCRIDSSNPGLQTSLNTVNTTEHVSVITVD | 730 | Db | 121 | VG1ADSYTNWFGPESOLVFLDKFILNGAGNWLAQOTRKHRPKDGMPVSTAQRWSTLH | 180 |
| Db | 661 | WTQFVNNTFCOMEPLITRTRIAVYFQPYINVSSCRIDSSNPGLQTSLNTVNTTEHVSVITVD | 720 | Qy | 397 | TEVIWYDQLTOPPADYGTAKHTFPNWGVITYGALPNTOQNTFYSFKSGKLGGGRAYV | 456 |
| Qy | 731 | YHNLKTRFNYLGPGFASTADQGQITRFGLTQIKVPRHDRLIFPQFKENIAGVL | 788 | Db | 181 | TEVIWYDQLTOPPADYGTAKHTFPNWGVITYGALPNTOQNTFYSFKSGKLGGGRAYV | 240 |
| Db | 721 | YHNLKTRFNYLGPGFASTADQGQITRFGLTQIKVPRHDRLIFPQFKENIAGVL | 778 | Qy | 457 | DIVHFQPSWIDGWSRSTNPQGHEDPDQNSFTPAFNGQFVSEALYGPPLSHLNNVLFAPS | 516 |
| | | | | Db | 241 | DIVHFQPSWIDGWSRSTNPQGHEDPDQNSFTFAEKGQFVSEALYGPPLSHLNNVLFAPS | 300 |
| <hr/> | | | | | | | |
| RESULT 3 | | | | | | | |
| Q9HAKS | | PRELIMINARY; | | PRT; | 755 AA. | | |
| AC | Q9HAKS; | | | | | | |
| DT | 01-MAR-2001 | (TREMBrel. 16, Created) | | | | | |
| DT | 01-MAR-2002 | (TREMBrel. 22, Last annotation update) | | | | | |
| DT | 01-OCT-2002 | (TREMBrel. 22, Last annotation update) | | | | | |
| DE | Hypothetical protein FLJ11477. | | | | | | |
| OS | Homo sapiens (Human). | | | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo; | | | | | | |
| OX | NCB_ TaxID=9606; | | | | | | |
| RN | [1] | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | |
| RC | TISSUE=Whole embryo; | | | | | | |
| | PubMed=14702039; DOI=10.1038/ng1285; | | | | | | |
| RA | Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., | | | | | | |
| RA | Watamatsu A., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., | | | | | | |
| RA | Seike M., Obayashi M., Nishii T., Shishibahara T., Tanaka T., Ishii S., | | | | | | |
| RA | Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., | | | | | | |
| RA | Murakami K., Yasuda T., Iwayanagi T., Waga H., Kondo H., Shiratori A., | | | | | | |
| RA | Sudo H., Hosogiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., | | | | | | |
| RA | Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., | | | | | | |
| RA | Abe K., Kamiya K., Matsuta N., Sato K., Tanikawa M., Yamazaki M., | | | | | | |
| RA | Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Murakawa K., | | | | | | |
| RA | Tanai H., Kimata M., Watanabe M., Hirzuka S., Chiba Y., Ishida S., | | | | | | |
| RA | Ono Y., Takiguchi S., Watanabe S., Yosida T., Horita T., Kusano J., | | | | | | |
| RA | Kaneko H., Watanabe A., Watanabe T., Sugiyama A., Takekoto M., Kawakami B., | | | | | | |
| RA | Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., | | | | | | |
| RA | Masahiko K., Komiyama M., Tashiro H., Oshima A., Sasou S., | | | | | | |
| RA | Yoshihiko Y., Yuuki H., Oshima N., Shiohata N., Sano S., | | | | | | |
| RA | Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O., | | | | | | |
| RA | RA Hishigaki H., Watanabe A., Mizoguchi H., Senoh A., Shigeoka S., | | | | | | |
| RA | RA Hishigaki H., Watanabe A., Sugiyama A., Takekoto M., Kawakami B., | | | | | | |
| RA | RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., | | | | | | |
| RA | RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikeda Y., Okamoto S., | | | | | | |
| RA | RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Shigeoka S., Senoh A., | | | | | | |
| RA | RA Matsumura K., Nakanuma Y., Mizuno T., Morinaga M., Sasaki M., | | | | | | |
| RA | RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., | | | | | | |
| RA | RA Mizushima-Sugano J., Saito T., Shirai Y., Takahashi Y., Nakagawa K., | | | | | | |
| RA | RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., | | | | | | |
| RA | RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S., | | | | | | |
| RT | "Complete sequencing and characterization of 21,243 full-length human | | | | | | |
| RT | CDNAs." | | | | | | |
| RL | Nat. Genet. 36:40-45 (2004). | | | | | | |
| DR | EMBL; AK021539; BAB13840_1; -; A8104C69FA4D6250 CRC64; | | | | | | |
| SEQUENCE | 755 AA; | 86162 MW; | | | | | |
| Qy | 61-84; | Score 4037; | DB 2; | Length | 755; | | |
| Best Local Similarity | 99.5%; | Pred. No. | 2..4..e-284; | | | | |
| Matches | 750; | Conservative | 2; | Mismatches | 0; | Gaps | 0; |
| | | | | | | | |
| Qy | 217 | MYEYSKVRSWKGOLLHNHOATNMIALTGALVTGYDKGSKANIKWQAVYDVKEMTMELLN | 276 | | | | |
| Db | 1 | MYEYSKVRSWKGOLLHNHOATNMIALTGALVTGYDKGSKANIKWQAVYDVKEMTMELLN | 60 | | | | |
| Qy | 277 | HIVDSSLDEGVAGSYTAKSVTQYVFLAQRFHNINNLDDNNWLKMHFWFYATLPGFORT | 336 | | | | |
| <hr/> | | | | | | | |
| RN | SEQUENCE FROM N.A. | | | | | | |
| RP | STRAIN=C57BL/6J; | TISSUE=Diencephalon; | | | | | |
| RC | DT 01-MAR-2003 (TREMBrel. 23, Last sequence update) | DT 01-MAR-2003 (TREMBrel. 23, Last annotation update) | | | | | |
| RX | MEDLINE=99279253; | | | | | | |
| RA | Carninci P., Hayashizaki Y., | | | | | | |
| RA | "High-efficiency full-length cDNA cloning."; | | | | | | |
| RA | DE CARCINOVA ANTIGEN RECOGNIZED BY T CELL (Fragment). | | | | | | |
| RA | DE Library, clone: 9330132809 Product: weakly similar to SQUamous CELL | | | | | | |
| RA | DE Name: 9330132809R1K. | | | | | | |
| RA | OS Mus musculus (Mouse). | | | | | | |
| OC | OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | | | |
| OC | NCBI_TaxID=10090; | | | | | | |
| RN | [1] | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | |
| RC | STRAIN=C57BL/6J; | TISSUE=Diencephalon; | | | | | |
| RX | DT 01-JUN-2003 (TREMBrel. 24, Last annotation update) | DT 01-JUN-2003 (TREMBrel. 24, Last annotation update) | | | | | |
| RA | Mus musculus (Mus musculus) male diencephalon cDNA, RIKEN full-length enriched library, clone: 9330132809 Product: weakly similar to SQUamous CELL | | | | | | |
| RA | DE CARCINOVA ANTIGEN RECOGNIZED BY T CELL (Fragment). | | | | | | |
| RA | DE Library, clone: 9330132809R1K. | | | | | | |
| RA | OS Mus musculus (Mouse). | | | | | | |
| OC | OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | | | |
| OC | NCBI_TaxID=10090; | | | | | | |
| RN | [1] | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | |
| RC | STRAIN=C57BL/6J; | TISSUE=Diencephalon; | | | | | |
| RX | DT 01-MAR-2003 (TREMBrel. 23, Last sequence update) | DT 01-MAR-2003 (TREMBrel. 23, Last annotation update) | | | | | |
| RA | MEDLINE=99279253; | | | | | | |
| RA | Carninci P., Hayashizaki Y., | | | | | | |
| RA | "High-efficiency full-length cDNA cloning."; | | | | | | |
| RA | DE Meth. Enzymol. 303:19-44 (1999). | | | | | | |
| RA | [2] | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | |
| RC | STRAIN=C57BL/6J; | TISSUE=Diencephalon; | | | | | |

| | |
|--|--|
| RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500; | Db 176 EDFLYNLGNGRKQKYLEKIVTEEMEYSSKIRSGWKGQLHNHDTATNMALLIGALVTG 235 |
| RA "Functional annotation of a full-length mouse cDNA collection.", Nature 409: 685-690 (2001). [3] | Oy 251 VDGSKANIWQAVWDMEKIMFLNHWVGSLDGVAWSYTAKSVTQYFLAQHFN 310 |
| RN SEQUENCE FROM N.A. | Db 236 VDGSKANIWQVQVVDMEKTMFLKHIVDGLDEGVAYGSYSKSVTQYFLAQHFN 295 |
| RC STRAIN=C57BL/6J; TISSUE=Diencephalon; | Oy 311 NNLDNNWLNKHEFWYATLILPGFORTYGIADSNTWFGPESOLVLDKFILKNGAGNW 370 |
| RA THE RIKEN Genome Exploration Research Group Phase I & II Team; | Db 296 NNFDNNWLNKHEFWYATLILPGFORTYGIADSNTWFGPESOLVLDKFILKNGAGNW 355 |
| RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.", Nature 420:563-573 (2002). [4] | Oy 371 AQQIRKRPKDGPMPYSTAQWSTLHTEYIWDPQTPOPPADYGTAKIHTPNNGVTVY 430 |
| RN SEQUENCE FROM N.A. | Db 356 AQQIRKRPKDGPMPYSTAQWSTLHTEYIWDPQTPOPPADYGTAKIHTPNNGVTVY 415 |
| RP SPRAIN=C57BL/6J; TISSUE=Diencephalon; MEDLINE=049374; PubMed=11042159; DOI=10.1101/gr.145100; | Oy 431 GAGLPNTQTNTFSFKSGKLGRAYDVIFHQPSMIDGWRSFNPQCHEPDQNSTFAPN 490 |
| RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konn H., Okazaki Y., Muramatsu M., Hayashizaki Y.; "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.", Genome Res. 10:1657-1650 (2000). [5] | Db 416 GGGLPNTQTNTFSFKSGKLGRAYDVIFHQPSMIDGWRSFNPQCHEPDQNSTFAPN 475 |
| RN SEQUENCE FROM N.A. | Oy 491 GQFVSEALYGPKLSHLNNLVFAFSPSSOCNKPKMEQOLGBCAOWLKWTCBQWLNKWTGEVGDAE 550 |
| RC SPRAIN=C57BL/6J; TISSUE=Diencephalon; MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600; | Db 476 GQFVSEALYGPKLSHLNNLVFAFSPSSOCNKPKMEQOLGBCAOWLKWTCBQWLNKWTGEVGDAE 535 |
| RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi Y., Kitaunishi T., Tashiro H., Ichijo M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishihira T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujikawa S., Inoue K., Togawa Y., Iwasa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuraya S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.; "RIKEN integrated sequence analysis (RISA) system-B4-format sequencing pipeline with 384 multicapillary sequencer.", Genome Res. 10:1757-1771 (2000). [6] | Oy 551 ITASQHGMVFSGEAVSAYSSAMRLKSYTRALLINQSOTLVVDDHIERQEDSPINSVA 610 |
| RP SEQUENCE FROM N.A. | Db 536 ITAAQHGMVFSGEAVSAYSSAMRLKSYTRALLINQSOTLVVDDHIERQEDSPINSVA 595 |
| RC SPRAIN=C57BL/6J; TISSUE=Diencephalon; MEDLINE=049374; PubMed=11042159; | Oy 611 FF 612 |
| RA Fukuda S., Puruno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayashi F., Hiramoto K., Hiraoka T., Hirozane T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.; Submitted to the EMBL/GenBank/DBJ databases. | Db 596 FF 597 |
| RA GO:GO:005615: C:extracellular space; TAS. | RESULT 5 |
| FT NON-TER 597 | O9H951 PRELIMINARY; ID O9H951; AC O9H951; DT 01-MAR-2001 (TREMBLrel. 16, Created); DT 01-OCT-2001 (TREMBLrel. 22, Last sequence update); DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update) |
| SQ SEQUENCE 597 AA; 68197 MW; E344F33EB16EDDC CRC64; | DB Hypothetical protein FLJ33005. |
| RA | OS Homo sapiens (Human); |
| RA | OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| RA | ONCBI_TaxID=9605; RN [1]; RP SEQUENCE FROM N.A. |
| RA | RR Pubmed=14702039; DOI=10.1038/ng1285; RX Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., RA Wakanatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., RA Seiken M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., RA Sudo H., Hosoi T., Kaku Y., Koda R., Sugawara M., RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., RA Ninomiya Y., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., RA Tanai H., Kimata M., Watanabe M., Hiracka S., Chiba Y., Ishida S., RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Horuta T., Kusano J., RA Kanehori K., Takahashi-Puji A., Hara H., Tanase T., Nomura Y., RA Togia S., Konai F., Hara R., Takeuchi K., Arita M., Imose N., RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., RA Yoshihikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., RA Hishigaki H., Watanabe T., Sugiyama A., Takekoto M., Kawakami B., RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., RA Fujimori Y., Komiyana M., Tashiro H., Tanigami A., Fujiwara T., RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., RA Kawabata A., Hikiji K., Kobatake N., Inagaki H., Ikeda Y., Okamoto S., RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., |
| RA | Qy 11 MAMFTGHLLFLMMAFSTFEESVNTSEWAFTDDQFTKQYDPRNQKLKKSM 70 |
| RA | Db 1 MAMFTEHFLFLMMCSFSTCBVSNSYSEWAFTDDQFTKQYDPRNQKLKKSM ---- 55 |
| RA | Qy 71 LHPFLPDAGEIQAMROKSRASHLFAIRSAVTVLNSPNTYLPPLPKHADFAAKWNEI 130 |
| RA | Db 56 LHPFLPDAGEIQAMROKSRASHLFAIRSAVTVLNSPNTYLPPLPKHADFAAKWNEI 115 |
| RA | Qy 131 YGNLNLPLPLAYCLCPDKVAFPFVLEYMDRMYGYKDWLVEAPGDEVPIGHSLTGFATA 190 |
| RA | Db 116 YGNLNLPLPLAYCLCPDKVAFPFVLEYMDRMYGYKDWLVEAPGDEVPIGHSLTGFATA 175 |
| RA | Qy 191 EDFLYNLGNGRKQKYLEKIVTEEMEYSSKIRSGWKGQLHNHOATNMALLGALVTG 250 |

| | | |
|----------|---|---|
| RA | Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai T., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.; "Complete sequencing and characterization of 21,243 full-length human cDNAs"; RT | "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001). [3] |
| RL | DR | SEQUENCE FROM N.A. |
| SQ | SEQUENCE 480 AA; 54559 MN; AFF24651B83E9DB CRC64; | SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; The FANTOM Consortium; the RIKEN Genome Exploration Research Group Phase I & II Team; RT |
| QY | Query Match 39.8%; Score 2597; DB 2; Length 480; Best Local Similarity 99.8%; Pred. No. 5.9e-180; Matches 479; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."; Nature 420:563-573 (2002). [4] |
| Db | 1 MHEFVYYATLLPQFORTVGIADSNNWYFGPEQSOLVEDKFLKGAGNWLAQIQRKHP 379 | RL SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100; |
| Db | 2 XDGPMVPSTAQRWSTLHTEIWYDQLTPPPADYGTAKHTFPNGVTVYGLPLNTOT 439 | RC Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.; RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630 (2000). [5] |
| QY | 440 NTFYSFKSGKLGRGAVYDTHFOPSYWIDGWRSPNPGHEPDONISFTFAPNGQFVSEAL 499 | RL SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; MEDLINE=10350913; PubMed=11076861; DOI=10.1101/gr.152600; |
| Db | 121 NTFYSFKSGKLGRGAVYDTHFOPSYWIDGWRSPNPGHEPDONISFTFAPNGQFVSEAL 180 | RC Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Fukuda S., Puruno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T., Hara S., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T., Katoch H., Kawai J., Koijima Y., Kondo H., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Obata N., Okazaki Y., Saitoh R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeeda Y., Tanaka T., Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases. [6] |
| QY | 500 YGPKLSHLNNVLFAPSPSSQCNPKEWGQGECAQWLKVTEGGDAEGBTTASQHGM 559 | RN SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; MEDLINE=10350913; PubMed=11076861; DOI=10.1101/gr.152600; |
| Db | 181 YGPKLSHLNNVLFAPSPSSQCNPKEWGQGECAQWLKVTEGGDAEGBTTASQHGM 240 | RC Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Fukuda S., Puruno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T., Hara S., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T., Katoch H., Kawai J., Koijima Y., Kondo H., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Obata N., Okazaki Y., Saitoh R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeeda Y., Tanaka T., Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases. [7] |
| QY | 560 VPFSCBAVSSAYSSAMRLKSYRALLNSCTLLVVDHIERQEDSPINSYSAFFNLDIF 619 | RN SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; MEDLINE=10350913; PubMed=11076861; DOI=10.1101/gr.152600; |
| Db | 241 VPFSCBAVSSAYSSAMRLKSYRALLNSCTLLVVDHIERQEDSPINSYSAFFNLDIF 300 | RC Adachi J., Aizawa K., Akimura T., Akimura T., Bono H., Carninci P., Fukuda S., Puruno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T., Hara S., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T., Katoch H., Kawai J., Koijima Y., Kondo H., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Obata N., Okazaki Y., Saitoh R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeeda Y., Tanaka T., Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases. [7] |
| QY | 620 KYIPKFMRHYANGAMMDVDAHYKMFWDHGNSPMASIQAEQAEEFKRWTQFVNTF 679 | RN SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; MEDLINE=10350913; PubMed=11076861; DOI=10.1101/gr.152600; |
| Db | 301 KYIPKFMRHYANGAMMDVDAHYKMFWDHGNSPMASIQAEQAEEFKRWTQFVNTF 360 | RC Adachi J., Aizawa K., Akimura T., Akimura T., Bono H., Carninci P., Fukuda S., Puruno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T., Hara S., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T., Katoch H., Kawai J., Koijima Y., Kondo H., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Obata N., Okazaki Y., Saitoh R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeeda Y., Tanaka T., Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases. [7] |
| QY | 680 QMESTITRAVYFYCPYINNSCRIPTDSSNPGLOTSIINYNTTERVSIIVDYNILKTRFN 739 | RN SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; MEDLINE=10350913; PubMed=11076861; DOI=10.1101/gr.152600; |
| Db | 361 QMESTITRAVYFYCPYINNSCRIPTDSSNPGLOTSIINYNTTERVSIIVDYNILKTRFN 420 | RC Adachi J., Aizawa K., Akimura T., Akimura T., Bono H., Carninci P., Fukuda S., Puruno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T., Hara S., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T., Katoch H., Kawai J., Koijima Y., Kondo H., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Obata N., Okazaki Y., Saitoh R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeeda Y., Tanaka T., Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases. [7] |
| QY | 740 YLGFGFFASVADQGQTTRFLGIGTQIVKPYTRHDLIFPFGEFKENITAVGLICSLVLTF 799 | RN SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; MEDLINE=10350913; PubMed=11076861; DOI=10.1101/gr.152600; |
| Db | 421 YLGFGFFASVADQGQTTRFLGIGTQIVKPYTRHDLIFPFGEFKENITAVGLICSLVLTF 480 | RC Adachi J., Aizawa K., Akimura T., Akimura T., Bono H., Carninci P., Fukuda S., Puruno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T., Hara S., Imotani K., Ishii Y., Itoh M., Kagawa T., Kasukawa T., Katoch H., Kawai J., Koijima Y., Kondo H., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Obata N., Okazaki Y., Saitoh R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeeda Y., Tanaka T., Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.; RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases. [7] |
| RESULT 6 | OSBB14 PRELIMINARY; PRT; 958 AA. | RN SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Brain; |
| AC | Q8BL14; 08BL14; 08BL14; | RC MEDLINE=7238257; PubMed=12477932; DOI=10.1073/dbna.24260899; |
| DT | 01-MAR-2003 (TREMbrel. 23, Created) | RC Klaunser R.D., Peingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Shanmnen C.M., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin B.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosesk S.A., McEwan P.J., McErnean K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muny D.M., Sodergren B.J., Lu X., Gibbs R.A., Fahey J., Heitton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shervchenko Y., Bourfard G.G., Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C., Butterfield Y.S., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schein J.E., Jones S.J., Marrs M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; RT RIKEN FANTOM Consortium; RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [8] |
| RN | 1 | RN SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| RC | SEQUENCE 10090; OS_musculus (Mouse); Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; NCBI_TaxID=10090; | RC MEDLINE=927953; PubMed=10319636; DOI=10.1016/S0076-6879(99)03004-9; |
| OC | 2 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OC | 3 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 4 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 5 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 6 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 7 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 8 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 9 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 10 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 11 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 12 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 13 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 14 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 15 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 16 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 17 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 18 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 19 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 20 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 21 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 22 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 23 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 24 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 25 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 26 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 27 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 28 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 29 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 30 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 31 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 32 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 33 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 34 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 35 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 36 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 37 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 38 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 39 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 40 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 41 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 42 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 43 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 44 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 45 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 46 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 47 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 48 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 49 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 50 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 51 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 52 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 53 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 54 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 55 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 56 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 57 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 58 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 59 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 60 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 61 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 62 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 63 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 64 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 65 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 66 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 67 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 68 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 69 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 70 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 71 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 72 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 73 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 74 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 75 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 76 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 77 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 78 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 79 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 80 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 81 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 82 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 83 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 84 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 85 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 86 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 87 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 88 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 89 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 90 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 91 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 92 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 93 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 94 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 95 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 96 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 97 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 98 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 99 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 100 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 101 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 102 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 103 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 104 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 105 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 106 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 107 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 108 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 109 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 110 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 111 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 112 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 113 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 114 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 115 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 116 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 117 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 118 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 119 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 120 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 121 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 122 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 123 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 124 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 125 | RC SEQUENCE FROM N.A. STRAIN=C57BL/6J; TISSUE=Parthenogenote; |
| OX | 126 | RC SEQUENCE FROM N.A. |

| | | | | | | |
|-----------------------|-----------|---|------------------------|------------------|---|-----|
| Db | 342 | LADQIRRNKVRBGGTPSKQRWCTLHTEFLWYDGLSKVSPPPDFGTPLTHYFBDWGVT | 401 | RT | CDNAS ⁿ | |
| RL | Nat. | Genet. | 36:40-45 (2004). | | | |
| ENBL | AK021849; | BAB1312.1; | - | | | |
| DR | 314 | 314 | | | | |
| FT | NON TER | 314 AA; | 35955 MW; | 722B5555DC065538 | CRC64; | |
| SQ | SEQUENCE | 314 AA; | | | | |
| Query Match | 20-41 | | Score 1331; DB 2; | Length 314; | | |
| Best Local Similarity | 32.41 | | Pred. 0; Mismatches 0; | | | |
| Matches | 313; | Conservative | | Indels 652; | Gaps 1; | |
| Db | 402 | YGSALPAEINRSLSFSFKSQRLLGGRAYDIDVHRNFKWDWKGMWNFGAHEHPDONSITFA | 461 | 1 | MPKGAPPWIMALMFTGHILLMLMAFPSTFEBSVSNYSEWAFTDDIDQFKTQKVQDF | 60 |
| Qy | 489 | PNGQVFVSEALYGPKLSHLNINVLPAPSQQCNKPWEQOLGE - CAQWLKWTGEEVGDA | 546 | 1 | MPKGAPPWIMALMFTGHILLMLMAFPSTFEBSVSNYSEWAFTDDIDQFKTQKVQDF | 60 |
| Db | 462 | PNGVPFITEALYGPKYTFNNVLMFSPAVSKCSPSPWVQVTEQCSKMSKVKHDLASC | 521 | 61 | RPNOKLKSMLHSLSLYFAGEIQAQRQSRASHILHLRAIRSANTVMSNPNTYLPPLPH | 120 |
| Qy | 547 | AGELITASQEGMVFVSGZAVSAYSSAMRKLSYSTRALLNSOTLVHDIERQEDSPIN | 606 | 61 | RPNOKLKSMLHSLSLYFAGEIQAQRQSRASHILHLRAIRSANTVMSNPNTYLPPLPH | 120 |
| Db | 522 | QGRVVAEEKNGVYVIRGBGVGAQNQLNKVNORNLLHPQLLVDQIHGEESPLE | 581 | 61 | RPNOKLKSMLHSLSLYFAGEIQAQRQSRASHILHLRAIRSANTVMSNPNTYLPPLPH | 120 |
| Qy | 607 | SVSFAFHNLIDDFKTYIPKFMNRTRGAMMDWDHYHKMFWFDHGNSPMSASIGEAQEAE | 666 | 61 | RPNOKLKSMLHSLSLYFAGEIQAQRQSRASHILHLRAIRSANTVMSNPNTYLPPLPH | 120 |
| Db | 582 | TAASPFHHNVDVDPFE ---ETVTDGTVHGAFTQRQDGlykonywMDDTGySEKAFASTVYPRG | 638 | 61 | RPNOKLKSMLHSLSLYFAGEIQAQRQSRASHILHLRAIRSANTVMSNPNTYLPPLPH | 120 |
| Qy | 667 | FKRWTQFVNNTFQMESTITRIATVYFQPYINVSSCRIDSSNPGQLQISLNVNNTHEVVS | 726 | 121 | ADFAAKKNEIYGNLNLPIPLALYCLCPEDKVAFFVLEYMDRMVGYKDWLIVENAPGDEVPI | 180 |
| Db | 639 | YPIYGNTNYNTVNTMLRSPITRAAYLFQISIDVQS-----PTVHGDQSQQLDVF | 686 | 121 | ADFAAKKNEIYGNLNLPIPLALYCLCPEDKVAFFVLEYMDRMVGYKDWLIVENAPGDEVPI | 180 |
| Qy | 727 | IVTDYHNLTFRNTYL----GFGCFASYV-ADGQI | 755 | 181 | GHSITGPATAFEDFLYNLLDNNHRRQKYLEKIKIWVITEEMEYSKVRSGKOLLHNHQATNM | 240 |
| Db | 687 | IATSKHAYAT---YLTWTGTGTQSAFQVADRHKI | 719 | 181 | GHSITGPATAFEDFLYNLLDNNHRRQKYLEKIKIWVITEEMEYSKVRSGKOLLHNHQATNM | 240 |
| Qy | | | | 241 | ALATGALVTVGDKGSKANIKWQAVVDDNEKTMFLNH.LVDGSLDEGVAGSYTAKSVTQY | 300 |
| Db | | | | 197 | | 196 |
| Qy | | | | 301 | VFLAQRFHFNINNLDNNNLMKMHMFYYATLPGFORTVGADSNNWYFGPESQLVFLDKF | 360 |
| Db | | | | 197 | | 196 |
| Qy | | | | 361 | ILXNGAGNMLAQQIRKRPKDGMVPSAQRWSTLHTEYIWYDQLTOPPADYGTAKIH | 420 |
| Db | | | | 197 | | 196 |
| Qy | | | | 421 | TFPNWGVVTFYAGLPLNQTNTFYVSFKSGKLGGRAVYDIVHFQPYSWIDGWSRNFPGHEHP | 480 |
| Db | | | | 197 | | 196 |
| Qy | | | | 481 | DQNSFTFAPNGQEVSEALYGPKLSHLNNVLFAPSPPSQNCNPWEGQLGECACQWLNWTG | 540 |
| Db | | | | 197 | | 196 |
| Qy | | | | 541 | EEVDAAGEBITASQHGMFVSGEAVASYSSAMRLKSYRALLLNSOTLLVVDHIERQ | 600 |
| Db | | | | 197 | | 196 |
| Qy | | | | 601 | EDSPINSVSAFFNLDIDPKYIPIVKMFRYNGAMMDVDAHYKNEWFDIHGNSPMASIQE | 660 |
| Db | | | | 197 | | 196 |
| Qy | | | | 721 | TEHVVISIVTDYHNLTFRNTYLFGCGFASVADGQQTREGTQAIKVPKVRHDLIFPFGF | 780 |
| Db | | | | 197 | | 196 |
| Qy | | | | 841 | WTRTEBAEGSKKSLSSBEGHMDLPDVVITSLPGSGAEILKQLFENSSDFLYIRVPTAYDI | 900 |
| Db | | | | 197 | | 196 |
| Qy | | | | 841 | WTRTEBAEGSKKSLSSBEGHMDLPDVVITSLPGSGAEILKQLFENSSDFLYIRVPTAYDI | 900 |
| Db | | | | 197 | | 196 |
| Qy | | | | 781 | KFNIAVGLLICISAVLTQWRFVYLSFRKLMRWILLYIALWFIELDWSTCSQPICAK | 840 |
| Db | | | | 197 | | 196 |
| Qy | | | | 841 | WTRTEBAEGSKKSLSSBEGHMDLPDVVITSLPGSGAEILKQLFENSSDFLYIRVPTAYDI | 900 |
| Db | | | | 197 | | 196 |
| Qy | | | | 901 | PETELEIOSFDACEWKYSDIRSCHFRILRGMLSIVQPTKLHLQNTHHEPNRGKLAQY | 960 |
| Db | | | | 249 | PETELEIOSFDACEWKYSDIRSCHFRILRGMLSIVQPTKLHLQNTHHEPNRGKLAQY | 308 |
| RT | | "Complete sequencing and characterization of 21,243 full-length human | | | | |

| | | | |
|--|--|--------------|--|
| Qy | 254 GSKANIWKOAAXVDYMEKTMFLNNHIVGDSLDEGVAYGSYTAKSVTQVYFLAQPH----- | 307 | DR EMBL; AE009232; AAL41841.1; -. |
| Db | 171 GRSE -WTKARANLETVLTMPLPE -DGSSEGTVYWRGYCPVPLAIAHTDVLVRQEADLW 226 | | DR EMBL; AE008581; AAK90358.1; -. |
| Qy | 308 ----FNNINLDDNNWLKGMHFWFYYATLPLPGFQRTVGLIADSNSYNNFYGPESQLVPLDFKLK 363 | | DR PIR; AC29228. |
| Db | 227 STGGFLRNT -TRN -RLH -----QSAPSPFEENIDHGDC -HDRLSGHSVALYVRLASAYQ 276 | | DR GO; GO:0016329; F:lyase activity; IEA. |
| Qy | 364 NGAGNWLAQCI -RGH-----RPDKGPMVPTSAQRWSLSTLHTEYIWYDQLTP-QPP 411 | | KW Lyase; Complete proteome. |
| Db | 277 DGTQMLGNLVAERHFWREAYESGVRPGYMP-----EAFLLEIWYDPRVTPAAPD 326 | | SEQUENCE 776 AA; 87870 MW; F588105922563E3 CRC64; |
| Qy | 412 ADYGTAKHTFPNGVVTTVAGLENPTQINTFVSPFKSGKLGKGRAYDIVHQPYSWIDOWR 471 | | Query Match 2.9%; Score 192; DB 2; Length 776; |
| Db | 327 REPCTA -YFPDLGQITARTGNDAA --TCVSKKAAPCGGHRRAWDECH -RLKAAGGW 379 | | Best Local Similarity 19.1%; Pred. No. 8e-05; |
| Qy | 472 SFPNCHEHPDQNNSFTFAPNGQVFVSEALYLG --PKLSHLNNVLYFAPSPSSQCNKPWEQOL 529 | | Matches 137; Conservative 94; MisMatches 279; Indels 208; Gaps 31; |
| Db | 380 AMSAHHHPDAGAFVLSHSGAFLAVDEGSNHKAHHNLVLV-----DGEGWA 428 | | Qy 33 EESVSNSYSEAWATDQDIDQFKTQKVODFRPNOKLKKSM-----HPSLYFDA 79 |
| Qy | 530 GECAQWLKWTG -TEVGDAEGETATASOFGEMTFVSGEAVSSAMPLKSYVRAILLN 587 | | Db 99 QKSATAHNST-----VRSPISEALPKTPLGRSARHAAQTSHERLWLN 146 |
| Db | 429 DE-GRYHYEGIPPERARRVARDVLI-AQPG-FAHATAESAAMFSERLGYQVRDRTLVYTP 484 | | Qy 80 GEIQAMRQ -KSRASHLHLHFRARSATVMSNPETYUPLL -PKHADFAKKNEEYGNNL 135 |
| Qy | 588 SQTLUVVDHIERQ 600 | | Db 147 EQSAPAFADAVAKDOPNHCWGAEFYEKSTEPWLRLPVMPEQPYDNNTVTALWROMY --- 202 |
| Db | 485 LGRVVILDELEAE 497 | | Qy 136 PPLALYCCLCPCDKVAFBVELEYMDRIVGYKD-----WLV -----BNAPGDE 177 |
| Qy | | | Db 203 -----IDCQEVIYAIRH-LATAGRVLGRDLDLDAASKWLLVAAWDTKGATSRAYNDE 254 |
| Qy | | | Qy 178 VPIGHSLTGTFATADFELYNLL-----DNHR --ROKYLEKIVWITEEMEYSKVRSGKQLLHN 233 |
| Db | | | Db 255 AGF -RVVIALANGYDWLDFHLSSEDRTRVSLBERTREVADIVIAHARIHVFP --YDS 310 |
| Qy | | | Qy 234 HQATNMIAALLTGAUTGVYDGSKSKANIKWAOAVVDMETKTMFLNNHIVDSLDE-----G 286 |
| Db | | | Db 311 HAVERSIAVLTPCAIALQGESDAEGEWLDYTVEFL -ATLYSPWAGTDGSAEGPHWMTC 369 |
| Qy | | | Qy 287 VAYGSYTAKSVTQVY -FLAQRFHNINNLDNNWLMKMHFVYATALLPQFQRTVQJADSNY 344 |
| Db | | | Db 370 MAYLIEANLIRSYIGDLYQRPF -----PONTGRCFPLYTKAPSTRAN - 413 |
| Qy | | | Qy 345 NWFPYGPESOLVLFDKF -----ILKNGAGNWLAQOIRKHRPKDGPMPVPTAQWMS 393 |
| Db | | | Db 414 ---FGDGSTLGDPGLKUGYNRQFAGTGNGHXYQWYDPIHK -----ADATGTE 459 |
| Qy | | | Qy 394 TLHTBYIWTD -----POLTOPPADYGTAKIHTPPNWGVVITYGAGLPNTQTNT 441 |
| Db | | | Db 460 MAFNYGWDLNFDLWYRHDYQVEAVSPADLPALEYDDIGMATIONQKDMDPDRHLQ - 518 |
| Qy | | | Qy 442 FVSFKSGKLUKGGRAYDVTHFQPSWIDGRSFENGFHEHPDQNSPTFAPNGQVFVSEALYK 501 |
| Db | | | Db 519 FV-RKS -----SPYGSLL -----SHSHGDQNAFVLFYAHGEDLAIQSGY - 554 |
| RESULT 12 | | | Qy 555 -----YVAF -----NSOMHLLAWRRQTRSNAVLIGGKQYAEMKDALARRAAGRIVSV 602 |
| QSUBJ1 | ID QSUBJ1 PRELIMINARY; | PRT; 776 AA. | Qy 554 SQHCEMFVSGEAVSSAMRL KSVYRALLLNSQTLVVDHIERQE ----- 601 |
| AC | QSUBJ1 ; Q7CRE1; | | Db 603 BEQPGHWRIVGDATAAYQVANPLVQKLRETHFVNDSYEVIVECECSRQELQMLCHTL 662 |
| DT | 01-JUN-2002 (TREMBLrel. 21, Created) | | Qy 602 DSPINSVS AFFNDIDFKYKIPYKMMNRYNGAMDWDYAHKNWFDHGNSPMASIQ 659 |
| DT | 01-JUN-2002 (TREMBLrel. 21, Last sequence update) | | Db 663 GAPQTGRSSF -----RYNGRKA --- 699 |
| DT | 25-OCT-2002 (TREMBLrel. 28, Last annotation update) | | |
| DE | 1190 ligninate lyase (AGR L 3558p). | | |
| GN | OrderendocuName=AGR L 3558; AtU3025; | | |
| OS | Agrobacterium tumefaciens (strain C58 / ATCC 33970). | | |
| OC | Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; | | |
| OC | Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium. | | |
| NCBI_TaxID | 176299; | | |
| RN | "The genome of the natural genetic engineer Agrobacterium tumefaciens | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=DUPONT; | | |
| RX | MEDLINE=1608550; PubMed=11743193; DOI=10.1126/science.1066804; | | |
| RA | Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr., Chapman P., Clendinnen J., Deatherage G., Gillett W., Grant C., Kurtykav T., Levy R., Li M.-J., McClelland B., Palmieri A., Raymond T., Rouse P., Saenphimmachak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W., | | |
| RT | "The genome of the natural genetic engineer Agrobacterium tumefaciens | | |
| RT | [2] | | |
| RN | Science 294:2317-2323 (2001). | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=Cereon; | | |
| RX | MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803; | | |
| RA | Goodner B., Hankin G., Gattung S., Blanchard M., Ourolo B., Goldman B.S., Cao Y., Askerazi M., Halling C., Mullin L., Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F., Wolam C., Allinger M., Dougherty D., Lapas C., Markez B., Flanagan C., Crowell C., Lemo C., Sear C., Strub G., Cielo C., Slater S., | | |
| RA | "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58."; | | |
| RT | Science 294:2323-2328(2001). | | |
| RL | | | |
| RESULT 13 | | | |
| Q9EC00 | PRELIMINARY; | PRT; 411 AA. | |
| ID Q9EC00 | PRELIMINARY; | PRT; 411 AA. | |
| AC Q9EC0; | | | |
| DT 01-MAR-2001 (TREMBLrel. 16, Created) | | | |
| DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update) | | | |
| DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update) | | | |
| DE Galactose 6-O-sulfotransferase GST-1 (Carbohydrate Gal-6) sulfotransferase 1. | | | |
| DE Name=Chaeli; Synonyms=Gct1; | | | |
| OS Mus musculus (Mouse) | | | |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelostomi; | | | |

Search completed: June 23, 2005, 08:49:05
 Job time : 138.102 secs

| | | | | |
|----|------|--|-----------------------------|------|
| Qy | 233 | NHQATNMIAILTGALVTGYDKGSKANIWKQAVVDDVMEMKTMLFLNHNIVDGSLLDEGVAYGSY | 292 | |
| Db | 1213 | -----GLYCTGQ----- | -----BETLJMYFYSQSS-----YSSY | 1239 |
| Qy | 293 | T---AKSYTQQVFLAQRFHENINLD-----NNWLKHWFYWFTATLPGFORTVGIAADSNNWF | 347 | |
| Db | 1240 | TDNNAPVTGLYIADMSDDMMTAQATNYWNNSY-----PQFDTVMADDSD----- | 1286 | |
| Qy | 348 | YGPESQLVPLDKEFLKGAGNWLAQQTIRKHPKGPMYSTAQWSTLHTHEYIWYDQLT | 407 | |
| Db | 1287 | -----NKKVITRVRNNRYAEDYE-----IPSSV-----TSNSNYSWGDSHLT | 1323 | |
| Qy | 408 | PQPADYGTAIKHTPPNGVWTVYAGLN-----TQTNFTVS-FKSGKLGRBAV | 455 | |
| Db | 1324 | -----MJJGGSVPNITFSEAAELRLSSTNMAISIIHNGYAGTRRI | 1363 | |
| Qy | 456 | YD1VHFOPSYSWIDGWRSFNPGHFPDONSFTFAPNGQVFVSEALYGPKLSHUNNVLYFAP | 515 | |
| Db | 1364 | QCNLMKQYASLGDKFIITDSFDDANFRNLP-----IF-----KFGKDENSDDSICYNE | 1414 | |
| Qy | 516 | SPSSOCNPKWPQEGOLGECAQWLKNTGE-----EVNSDAEGLITASQHGMVFMVSEAVSAYS | 571 | |
| Db | 1415 | NPSSSE-DCKMFSSKDDNKTADNGGTOCI DACTSNKDFYNIQEITAVSTGGYNSSYK | 1473 | |
| Qy | 572 | SAMBLKSVTYRALLILNSQTLVVDHIERQDSPINSVSASFNLIDFEKYIPIKFENRYN | 631 | |
| Db | 1474 | IS-----NPNNI-----NTGIDSAYV-KVTVKAG | 1496 | |
| Qy | 632 | GAMMDVWDHYRMFWFDHHGNSPMASIOEAQQAEFGRWTQFVNTEQMEST--ITRIA | 689 | |
| Db | 1497 | G-----DQOIFTADISTYVFP-----QPAPSFEEMIQFENNLTIDCKNLNFIDNOA | 1542 | |
| Qy | 690 | YVFYGPYINSSCRFDSSNPGLOISLNUVNTHEVSVTUVTHYNL-----KTRFENYL | 741 | |
| Db | 1543 | HIEIDFTAQDGRFLGAETFTIPVTKVLGTENVIALYSEENGVQMIGAYRTRANTL | 1602 | |
| Qy | 742 | GFGGFASVADQ-----QITRFLGFOAIVKPVHRDRIIPFGFKENIANGLI | 789 | |
| Db | 1603 | FAQQLVSRANGIDAVLSMETONIQEPOLGAQGYVYVLDKYDESHGTNKSFAIEYVDI | 1662 | |
| Qy | 790 | LCISLIVILTFQWRYLSPEKLMWILILVIALMFIELL----DVAWSTCQSOPICAKW----- | 841 | |
| Db | 1663 | FKENDSFVYIYQGESETOTVVK----VFLSYSPFIAETGNKRHLW----VRAKYQKET | 1711 | |
| Qy | 842 | -----TRTEAE-----GSKKSLSSSEGFGHMDLPDVVITSULPGSGA----- | 875 | |
| Db | 1712 | TDKLFDRDTEKOPHGWFLSDDHTKTFSGLSISAGALKNDSEMPD-----SGNAL | 1761 | |
| Qy | 876 | -----EILKQFLFNSSD-----FLYCRVPTAYIDIPETELEIDSVDACEWKV | 918 | |
| Db | 1762 | YFWELFYXTPMMMMAHRLIQEONPDAANHFWRYWSPSCYI-----VDGKIAITYHWNV | 1813 | |
| Qy | 919 | SDRSRGHFLLRGMQLSLVQPTKUHLQNLHLPBN-----RGKLAQYFA-----MN | 964 | |
| Db | 1814 | -----RPLEEDTSWNAQQLDSTDPAVAQDDPMHYKUTAFATDLMAR | 1858 | |
| Qy | 965 | KDKKRKFKRRESLPEORSMQGAFDRDAYTRALR-----RHLVYYPSPARPVLSI-SSSG | 1018 | |
| Db | 1859 | GDAAYRQLERDTIABAKW-----YTOQANLNGDPOVMLSTWANTLGNAAK | 1908 | |
| Qy | 1019 | WTLKLHFFQEVIGASMRALYIVRDPRAMYISMLYNSKPSLXSLKNYVPEHLAKLFEKGGK | 1078 | |
| Db | 1909 | TTQOYR--OOVL-TOLRNURSVTP-----LIGTANSLTAL-FUQENSKL-----K | 1951 | |
| Qy | 1079 | G-----KCUNNSGYAFEEPEPLRKELSKSNAVSLSHMLANTAAARINTDILPT | 1130 | |
| Db | 1952 | GYWRTLQAORMENURHNLSDQPLSLPLYAKPADPKALLSAAVSASQGGA----- | 2001 | |
| Qy | 1131 | SYQLVKFEDIVH-FPQ | 1145 | |
| Db | 2002 | -DLPKAPLTIHFRFPQ | 2015 | |